

[illegible]

ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA
CCCACGCGTCCGGGCCGGGAGCAGCACGGCCGCAGGACCTGGAGCTCCGGCTGCGTCTTCCCCG
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG
CTGCTGCCGCCCGCGCCGGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA
CGGCTTGAGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCGCTGCTGGAGATC
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCTGAATGCAATCAGATGCTAGAGGCGCAGGAGGA
GCACCTGGAGGCCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTGAGTGGTTTT
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC
CAGGGCGGATCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA
GGGCGACGGGTCTTGCCGGTGCCACATGGGGTACCAGGGCCCGCTGTGCACTGACTGCATGG
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCCTCCCTGCAGCGCTG
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG
GGCTGCACAGGGGAAGGCCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG
AAAAGTGTCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTTGACGGCTTCGAAGAAACG
GAAGATGCCTGTGTGCCGCCGCGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT
GCCCTCCCGCGAAGACCTGTAATGTGCCGGACTTACCCTTTAAATTATTGAGAAGGATGTCC
CGTGGAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGTTTGTCTTA
AACAGACTTGTATATTTTGATACAGTTCCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC
TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA
TCACAAATTTACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTC
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC
ATCTCAATTAGTCAGCAACCCAGTTTTT

FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTL SKYESSEIRL
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFVCVTKLVCCSPGTYGPDCLACQGGSQRPCSG
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGECEVGWVLDE
GACVDVDECAAEPSPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAECT
CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343
and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and
313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

FIGURE 3

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACA
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG
GCACGAGGAGTTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC**ATG**CCCCGGAGGAGCGCCTTC
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC
CGGGCCGCCGAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA
TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCTGTCTTGGCGCTCCCTGGATA
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA
TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGAAAAACAGGATGGGGTGGCAGCATTGGAAGT
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT
TCTTTAAACATGTCAACAAGCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT
GAAAGACGCATCTGCGAGTGTCTGATGGGTTCACGGACCTCACTGTGAGAAAGCCCTTTG
TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG
GATTCTATGGAGTGAACGTGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC
TGTTTCTACCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG
CAAATGCCCAACCCCTGTGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGT
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT
GGAACCTGCCATGAACCCAAACAAATGCCAATGTCAAGAAGGTGGCATGGAAGACACTGCAA
TAAAAGGTACGAAGCCAGCCTCATACTGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG
TGAACTCCGACATCTGAAACGTTTTTAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCA
TGTGTTGAATGTTCAAATAATGTTCACTTACACTTAAGAATACTGGCCTGAATTTTATTAGCT
TCATTATAAATCACTGAGCTGATATTTACTCTTCTTTTAAGTTTTCTAAGTACGTCTGTAG
CATGATGGTATAGATTTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA
TCAGGTAAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT
GTCTGGGGGCAGGGGAACATCAGAAAGGTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATGTGTCAGATATTTAGAT
GTTTGTACATTTTTTAAAAATTGCTCTTAATTTTTTAACTCTCAATACAATATATTTTGACC
TTACCATTATTCCAGAGATTCACTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATT
AAACAATATAATATATTTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTGCAT
TGGCTTGAAGCAATATAATATATTGTAAACAAAAACAGCTCTTACCTAATAAACATTTTAT
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC
CGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC
RNGGFCNERRICECPDGFHGHPCHEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEEERDP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTTCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG
CCCCAGCCCACACCTTCACCAGGGCCCAGGAGCCACCATGTTGGCGATGTCCACTGGGGCTAC
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTTCGTGGGCGCCGG
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA
GCAGGACCTGTGCTGCCGCGGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC
TGCCTCGGCGTGCCACCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA
TCCAGTCTTGGAACGTACTGGGACAACGTGAACCGTTGCACCTGCCAGGAGAACAGGCAGT
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG
GAACCACAGCGCCTTCTGGGGCATGACCCTGGATTGAGGGCATTCGCTACCGCCTGGGCACCA
TCCGCCCATCTTCCTCGGTTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA
CCAAGGCAACTGTGCAGGCTCCTGGGCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT
CAATCCATTCTCTGGGACACATGACGCCTGTCTGTGCGCCCAAGAACCTGCTGTCTTGTGAC
ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTTCG
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC
CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTTCGGGGCAAGCGCCAGGCCACTGCC
CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACCTCTGTCTACCGCCT
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA
TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC
CTTGGGAGGCCAGAGAGATAACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC
CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG
AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTTCATCACTGAGGCTG
CGGGCACACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCAATG
GGGCGGTGACCCACAGCCTCGCCCGACAGAGCCCGGGCGCAGGCGGGCGCCAGGGCGCTAAT
CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCTGGGAGCCGCGGGCAGGCGAGACTGGCG
GAGCCCCCAGACCTCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG
ATCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC
CCCAATACCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG
TTGCCCAGGTTGGAGTGAGTGAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA
AGTGACCCTCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACACACCTGGC
TAATTTTTGTATTTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTTGAACT
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC
ACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAA
TAAAACCAAAGTATTGATAAAAAAAA

FIGURE 6

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPPIQGCMHGGRIYPVLGTYWDNCNR
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,
411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,
269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site.

amino acids 398-409

FIGURE 7

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCCCCTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATAACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

Sequence for

FIGURE 8

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG
CAGCTTGAAGTGCCTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT
CTGCACCCCTGTCCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCCTCCAACCCTCTCTGCTGCTGTTTC
CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT
CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA
GGGGACAGGCACTCAGGAGGGCCCAGTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT
AATAAACACCTGTTGGATAAGCCAAAAAA

FIGURE 9

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT
LGVVPQASVPLLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHDPFPMALSRTPTRQISSDT
DPPADGPSNPLCCCFHGPFAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

FIGURE 10

CCCACGCGTCCGAACCTCTCCAGCGATGGGAGCCGCCCCGCTGCTGCCCAACCTCACTCTGT
GCTTACAGCTGCTGATTCTCTGCTGTCAAACCTCAGTACGTGAGGGACCAGGGCGCCATGACC
GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA
CGTGCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC
TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTTCGCATCAAAGGGGCTGAGAGTGAGAAG
TACATCTGTATGAACAAGAGGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG
CGTGTTACCGGAGATCGTGCTGGAGAACAACCTATACGGCCTTCCAGAACGCCCCGGCACGAGG
GCTGGTTTCATGGCCTTCACGCGGCAGGGGCGGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC
CAGCGCGAGGCCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCCACGC
CGAGAAGCAGAAGCAGTTCGAGTTTGTGGGCTCCGCCCCCACC CGCCGGACCAAGCGCACAC
GGCGGCCCCAGCCCCTCACGTAGTCTGGGAGGCAGGGGGCAGCAGCCCCCTGGGCGCCTCCC
CACCCCTTTCCTTCTTAATCCAAGGACTGGGCTGGGGTGGCGGGAGGGGAGCCAGATCCCC
GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCCAGCTGGGAAGGGGCAGGCCGGTG
CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT
GAGTGTACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC
CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA
TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACCTCCTCCTGGCTAGACTGTA
GGAAGGGACTTTTGTTTGTTTGTTTGTTTCAGGAAAAAAGAAAGGGAGAGAGAGGAAAATAG
AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCAACCCCAACTCCCAGCCC
CGGAATAAAACCATTTTCCTGC

FIGURE 11

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMATFTRQGRPRQASRSRQNRQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCCTCCTGTTGAATTTTTTGCACATGGAG
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
CATTTTGATTTTGCTGTTTATTTTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT
TTCCGTACTTCAGAAATGGGCCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCTT
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAACTCCTGGCCTGCC
CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC
TGGATTTCTTGCAAGACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC
AACTGGACGAATTCCTCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
AATATTCAGACCATTTCACGGGCTGCTCTTGGCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCTATATCCGACATGGCCTTCCAGAA
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTTCGCTGTCC
CACCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
ACTGCTCGGAATAACCTTGGTTTTTGTGACTGCAGTATTAAATGGGTACAGAAATGGCTCAA
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCTGAACAAGTCCGGG
GGATGGCCGTCAGGGAATTAATATGAATCTTTTGTCTGTCCACCACGACCCCCGGCCTG
CCTCTCTTACCCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCTCTCTAT
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAACTTCCCACGATTC
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCCTGATGGCATA
CAAACCTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTTCAGAGGC
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTCTGGCATATGCACAAAAGGGGCGCTACACCTC
CCAGAAGTGGAAATACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
AGGACAACCTCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT
CAACTCCTTAAAGGAGATTTTCACTGTCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTTGTAACCTTTTGCTTTTAAATCTT

FIGURE 13

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
EGVTVLVYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
SRAALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMVRELMNMLLSCPTTTPGLPLFTP
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGREVRTPPISERIQLSIHFVND
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
DAFNRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL
LSVFCWHMHKKGRYTSQKWYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

Year	Total population		Population aged 15 years and over		Population aged 65 years and over		Population aged 75 years and over	
	Number	Density	Number	Density	Number	Density	Number	Density
1980	1 000 000	100	600 000	60	100 000	10	20 000	2
1985	1 100 000	110	650 000	65	110 000	11	22 000	2.2
1990	1 200 000	120	700 000	70	120 000	12	24 000	2.4
1995	1 300 000	130	750 000	75	130 000	13	26 000	2.6
2000	1 400 000	140	800 000	80	140 000	14	28 000	2.8
2005	1 500 000	150	850 000	85	150 000	15	30 000	3.0
2010	1 600 000	160	900 000	90	160 000	16	32 000	3.2
2015	1 700 000	170	950 000	95	170 000	17	34 000	3.4
2020	1 800 000	180	1 000 000	100	180 000	18	36 000	3.6
2025	1 900 000	190	1 050 000	105	190 000	19	38 000	3.8
2030	2 000 000	200	1 100 000	110	200 000	20	40 000	4.0
2035	2 100 000	210	1 150 000	115	210 000	21	42 000	4.2
2040	2 200 000	220	1 200 000	120	220 000	22	44 000	4.4
2045	2 300 000	230	1 250 000	125	230 000	23	46 000	4.6
2050	2 400 000	240	1 300 000	130	240 000	24	48 000	4.8
2055	2 500 000	250	1 350 000	135	250 000	25	50 000	5.0
2060	2 600 000	260	1 400 000	140	260 000	26	52 000	5.2
2065	2 700 000	270	1 450 000	145	270 000	27	54 000	5.4
2070	2 800 000	280	1 500 000	150	280 000	28	56 000	5.6
2075	2 900 000	290	1 550 000	155	290 000	29	58 000	5.8
2080	3 000 000	300	1 600 000	160	300 000	30	60 000	6.0
2085	3 100 000	310	1 650 000	165	310 000	31	62 000	6.2
2090	3 200 000	320	1 700 000	170	320 000	32	64 000	6.4
2095	3 300 000	330	1 750 000	175	330 000	33	66 000	6.6
2100	3 400 000	340	1 800 000	180	340 000	34	68 000	6.8

ACTTTGGAGCAACGCGCGCGCGGAGACAGAGGCAGAGGCAGAGGCTGGGGCTCCGCTCCTCGCCTCCCAACGAGCC
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCTGCC
TCGCTTCCCAGGCGCGCGCGGCTGCAGCCTTGCCCCCTCTTGCTCGCCTTGAAAAATGGAAGAGATGCTCGCAGGCCT
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTACAGTGGGAGGTCCATCT
CTAGGGGCGAGACACGCTCGGACCCACCCGCGAGACGGCCCTTTCTGGAGAGTTCTCTGTGAGAACAAGCGGGCAGACC
TGGTTTTTCATCATTGACAGCTCTCGAGTGTCAACACCCATGACTGTCAAAAGTCAAGGAGTTTCATCGTGGACA
TCTTGCAATTCTTGGACATTGGTCTCTGATGTCAACGAGTGGGCTGTCTGCAAGATGAGGAGCATGTCAAGAATG
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTCTGCAAGAGGATGCGGCATCTGTCCACGG
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCGGCCCTGA
GGGAGAATGTGCCACGGGTCTAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA
AGGCACGGGACACGGGATCTTAATCTTTGGCATTTGGTGTGGGCGAGGTAGACTTCAACACCTTGAAGTCCATTG
GGAGTGAGCCCCATGAGGACCATTGTCTTCTTGTGGCAATTTTCAGCCAGATTGAGACGCTGACCTCCGTGTCTCC
AGAAGAAGTTGTGCACGGCCCAATGTGTCAGCACCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG
GCTCATACGTCTGCAGGTGCAACAAGGCTACATTTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT
GTGCCATGGAGACCACAACCTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGTCTACAGTG
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCCTCAGAAAACCACGGATGTGAAC
ATGATGTGTAAATGTGATGGCTCTTACCTTTGGCAGTGCCATGAAGGATTTGTCTTTAACCAGATGAAAAAA
CGTGCACAAGGATCAACTACTGTGCATGAACAAACGGGCTGTGAGCATGATGCGTCAACATGAGGAGAGCT
ACTACTGCCGTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAACCTGCAGCCGAGTGGACCACCTGTGCAC
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGTCTCAGAAGGCTTCC
TCTCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTACTGCCCTGCTGAGTGACCATGGTTGTGAATACTCCT
GTGTCAACATGGACAGATCTTTTGCTTGTCACTGTCTCGTGGGACAGCTGTCTCCGACGCGATGGGGAAGACGTGTG
CAAAATTTGGAATCTTTGTGCTCTGGGGTGACCAAGTTGTGAACATTCGTGTGAAGCAGTGAAGATTCGTTTGTGT
GCCAGTGTCTTTGAAGTTTATATACTCCGTGAAGATGGAAAACTTCGAGAAGGAAGATGTCTGCCAACGTATAG
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGTCTTGAGGGGATTCCCGCTCG
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAATCAACCCACCATGGCTGCGAACACATTTGTGTTA
ATAATGGGAATTCCTACATCTGCAATGCTCAGAGGGATTTGTCTAGCTGAGGACGGAAGACGGTGAAGAAAT
GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGAGAGAAGAGAATTTTGAGGTCTG
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTTGACAATTTCCCCAAAGCCGCTCGATGGGGCTGCTCCAGT
ATTCACACAGGTTCCACACAGAGTTCACTCTGAGAACTTCAACTCAGCCAAAGACATGAAAAAAGCCCTGCCCC
ACATGAAATACATGGGAAAGGGCTTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCCAAG
GAGAAGGGGCCAGGCCCCCTTCCACAAGGGTGCCCAAGCAGCCATTGTGTTCACCGACGGACGGGCTCAGGATG
ACGCTTCCGAGTGGGCGAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG
AGGAGGAACCTACAAGAGATTTCCCTCTGAGCCCAACAAACAGCATCTCTTCTATGCCGAAGACTTCAAGCAATGG
ATGAGATAAGTGAAAAACTCAAGAAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCAG
CAGGGGAACCTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTACCATAAAATATCCAAGACCTACTTT
CCTGTTCTAATTTTGCACTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTACGGTCTACACAAAAGCTTT
CCCATTCACAAAACCTTCAAGAACGCTTTTGAAGAAAAACAGCATCAATGCAAAATGTGAAAACCTTATAATGT
TCCGAACCTTGTCAAACGAAGAGTAAAGAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC
TGGAAAATCGCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCAATTGTATCAGGATTAACAT
GAACGCAGTGCAGAGCCCCAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA
GAAACCTGGTTTGCCAAGAACAAAGACAAGAAAGTATACATAACTTGTATAAATTTATCTAGGAAAAAAATCCT
TCAGAATCTTAAGATGAATTTACCAAGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAATATACTGTGGACAC
AACTTGCTTCTGCCTCATCTTCCCTTAGTGTGCAATCTCACTTTGACTATCAGATAAAGTTTGACAGTCTTACTT
CTGTAGAACACTGGCCATAGGAAATGCTGTTTCTTGTACTGGACTTACCTTGATATATGATATGATATGATG
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAATAATCACCACCTTCAG

FIGURE 15

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV
KMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKP GC
EHECVNMEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI
NEDLKTC SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTC AKLDSCALGDHGCE
HSCVSSSEDSFVCQC FEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA
EDGKRCRRKDVCKSTHHGCEHICVNNNGNSYICKCSEGFLAEDGRRCKKCTEGPIDLVFVID
GSKSLGEENFEVVKQFVTGIIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA
HMKYMGKGSMTGLALKHMFERSFTQEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN
GITMYAVGVGKAIEEELQEIA SEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS
PAGELPKTVQQPTSEPTINIQLDLSCSNFAVQHRYLFEEFNLLRSTQKLSHSTKPSGSPL
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,
781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,
639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,
540-546, 581-587

Parameter	Value	Parameter	Value
α	0.05	β	0.05
γ	0.05	δ	0.05
ϵ	0.05	ζ	0.05
η	0.05	θ	0.05
ι	0.05	κ	0.05
λ	0.05	μ	0.05
ν	0.05	ξ	0.05
\omicron	0.05	π	0.05
ρ	0.05	σ	0.05
τ	0.05	υ	0.05
ϕ	0.05	χ	0.05
ψ	0.05	ω	0.05
Ω	0.05	Θ	0.05
Φ	0.05	Ψ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05
Λ	0.05	Γ	0.05
Δ	0.05	Θ	0.05
Σ	0.05	Π	0.05
Υ	0.05	Λ	0.05
Γ	0.05	Δ	0.05
Θ	0.05	Σ	0.05
Π	0.05	Υ	0.05

GGAGCCGCCCTGGGTGTCTAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCTTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAAACAAGGCAAATCTAGGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCACCCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTATTGATGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS
SQPWEVPPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ
WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPPWKSSDTISKNGTL
SSVTSARALRPPHGP RP PGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSSGLSR
MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 18

bioRxiv preprint doi: <https://doi.org/10.1101/156666>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTTCCACTTTG
TTGAATTGTTTCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA
AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAA
TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC
ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA
AGACAGGTTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT
TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA
GAACCTGTGGCTTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA
TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA
CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT
AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC
ACATCTTACAAAACCTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC
AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT
TCATATAACATGAAACATATTCATCCTCATATGAATATGGATGGAGACTACATAAATATATT
TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA
AGAGTATTGGTCTTTTGCTTTTCATCATCTGACAACCTTCTATTGAAACCTCAAAATTATGAT
AATCTGAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTTCAGTCTCAATGAGCTCAAACCC
ACCCACATTATATGAACTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA
GGTATAGGAGTCTATGTGCATTTTGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT
TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT
GACACATTTTGCAATTTTGATGTCCTCTGGTCTTCCATTGGTATTAAAGATTATAATATTC
TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT
ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG
TAGCCTATTTCTTGCTGAACTTGTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT
TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTGCATGGATGTGC
ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCTCTACAACAAGGGATTTTTTGCA
CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC
TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACCTTTATT
TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTGGAGTCAT
CATATACAAAGTTTTTCGTACACTGCAGGGTTGAAACCAGAAGTTAGTTGCTTTGAGAACA
TAAGGTCTTGTCAGAGGAGCCCTCGCTCTTCTGTTCTTCTCGGCACCACCTGGATCTTT
GGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC
TTTCCAGGGGATGTTTATTTTTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT
ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGGTAAACATAGAGAATG
GTGGATAATTACAACCTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA
TGACTCATCAAATTATCCAATTATTAACACTAGACAAAAGTATTTTAAATCAGTTTTTCT
GTTTATGCTATAGGAACCTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT
TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTGGAAAGTAATTGGTTT
CTCAGGAGTGATATCACTGCACCCAAGGAAAGATTTTCTTCTAACACGAGAAGTATATGAA
TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTGCTTTGAACTAGTCC
CCTACCACCTCGGTAAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA
ATATCAAACAGTGAAAAGGGAATGATAAGATGTATTTTGAATGAACTGTTTTTCTGTAGAC
TAGCTGAGAAATTGTTGACATAAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA
TTGTTCTGAACTTAAATGTCCACTAAAACAACCTTAGACTTCTGTTTGCTAAATCTGTTTCTT
TTTCTAATATTCTAAAAAAGGTTTACCTCCACAAATTGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 19

MKRLPLLVFSTLLNCSYTONCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA
NINKTLTKIRSIKEPVALLOEVYRNSVTDLSPTDIITYIEILAESSLLGYKNNTISAKDTL
SNSTLTETFKTVNNFVQORDTFVVDKLSVNHRRLTKLMHTVEQATLRISQSFQKTTEFDT
NSTDIALKVFFFDSDYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS
SSDNFLLKPQNYDNSEEEERVISVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF
WNYSPTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSTIAGL
LHYFFLAFAWMCIIEGHLVIVGVYINKGFLHKNFYIFGYLSPAVVVGFSAAALGYRYYGT
TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA
LALLFLLGTTWIFGVLVHVASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV
PCCFGCLR

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,
648-664

N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,
188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,
346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,
434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

Case	Age	Sex	Duration	Location	Findings	Diagnosis	Outcome
1	25	M	10 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
2	35	F	5 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
3	45	M	2 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
4	55	F	15 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
5	65	M	8 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
6	75	F	3 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
7	85	M	12 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
8	95	F	7 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy

Case	Age	Sex	Duration	Location	Findings	Diagnosis	Outcome
1	25	M	10 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
2	35	F	5 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
3	45	M	2 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
4	55	F	15 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
5	65	M	8 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
6	75	F	3 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy
7	85	M	12 years	Left eye	Small, well-circumscribed, pigmented lesion	Benign melanocytic nevus	Stable
8	95	F	7 years	Right eye	Large, irregular, pigmented lesion with satellite lesions	Malignant melanoma	Enucleation, chemotherapy, and radiation therapy

[illegible][illegible]

FIGURE 22

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDP IRLRHLYTSGPHGLSSCFLRI
RADGVVDCARGQSAHSLLEIKAVLRITVAIKGVH SVRYLCMGADGKMQGLLQYSEEDCAFEE
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 23

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG
CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA
AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAAGTGTGGTAGAGCTACGATGTCAAGA
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATAACAATGAATACAAAACCTGGAAGT
CTGCAATTTAATACTGTTTCCAAACTGGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC
TGTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC
CACTGCAGTCCAGCCTGGGTAAACAGAGCAAGATTCCATCTCAAAAAATAAAATAAATA
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

FIGURE 24

MARRSRHRLLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE
WKKLGSRVSFVYYQOTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQQNLEED
TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST
NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNVSGYRRCPGKRMQVDDLNISGIIAAVVVVA
LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTVPVLPALWKAAGGSRGQEF

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 130-144, 238-258

N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

Amidation site.

amino acids 226-230

FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG
AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT
CATGCTGCTATTCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA
ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA
TTATATCATTAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTTTGGACAATG
CAATTGTGGCACTGGCACTTATTTTCAGTGAAGAAAAACTTTGTGGTTCTATGGCATTTCATCA
TTTGACAAATGCAAGCATCTTCTTATCAATCAGCTCCTATTGAACTTACTAGCACTGACTG
TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT
CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG
TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT
ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCAGCCAGATTGCC
AGCTAACACACAGATTCTTCTCCTACAGACTAACAAATATTGCAAAAATTGAATACTCCACAG
ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCAACCAAT
ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAACTTACTGA
ACTGCCTGAAAAATGTCTGTCCGAAGTGAAGCACTTACAAGAACTCTATATTAATCACAACCT
TGCTTTCTACAATTTACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC
AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT
TCTGATGATTGGGGAAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCA
ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT
GGACTGGAAAACCTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGT
TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTTGGATCTAAATAAAAAATCCTATTAATAGAA
TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAGAGTTGGGGATAAATAATATGCCT
GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC
TACTAACAAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTTCAGACTCCCCAAGCTGG
AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG
CCAAACCTCAAGGAAATCAGCATAACAGTAACCCCATCAGGTGTGACTGTGTATCCGTTG
GATGAACATGAACAAAACCAACATTCGATTTCATGGAGCCAGATTCAGTGTGTTTGCCTGGACC
CACCTGAATTCCAAGGTGAGAATGTTCCGCAAGTGCATTTTCAGGGACATGATGGAAATTTGT
CTCCCTCTTATAGCTCCTGAGAGCTTTCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT
TTCTTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG
GTCAAAAACCTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA
GATATAAATGGCGTAACCTCCCAAAGAAGGGGTTTATATACTTGTATAGCAACTAACCTAGT
TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG
GCTCTTTGAATATTAAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCTGGAAGCA
AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA
TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC
CATCAACTGAGTATAAAATTTGTATTGATATTTCCACCATCTATCAGAAAAACAGAAAAAA
TGTGTAAATGTCACCACCAAGGTTTGCACCTGATCAAAAAGAGTATGAAAAGAATAATAC
CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA
GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG
AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA
AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCTT
AAAAACCACCAAGGAAACCTACTCCAAAAATGAAC

FIGURE 26

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT
FPARLPANTQILLQLQTNNAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE
ENKLTPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLPD
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSPNPIRC
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGNVRQVHFRDMMEICLPLIAPESFPSNLNV
EAGSYVSFHCRTAEPPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTC
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE
YEKNNTTTLMACLGGLGIIGVICLISCLSPMNCDGGHSYVRNYLQKPTFALGELYPPLIN
LWEAGKEKSTSLKVKATVIGLPTNMS

Signal sequence:

amino acids 1-22

Transmembrane domain:

amino acids 633-650

N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,
608-612, 624-628, 625-629

Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

Tyrosine kinase phosphorylation site.

amino acids 570-579

N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,
491-497, 492-498, 634-640, 702-708

Cell attachment sequence.

amino acids 277-280

FIGURE 27

GCCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTAG
CTGCAGCCTTTTGAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA
TGATACTGTGCTTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT
GCCTTCAAAGGAGTAGCTGAAACCTTGACAGACTCTGGACTTGTCCGACAATCGGATTCAAAG
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCTCAATGCTGC
CAACGACGCTGACCTTTGTAACCTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA
TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGC AAAATCAGGAG
GATGCCCCGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAAACTGACTGTTCATTGAGAAAGAAAGAAA
GTAGTTTGC GATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA
ACTTTGTATTTCAGTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTT
AATTTAAAAGCAAATAAAAGCTTAACCTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAACA

FIGURE 28

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFLL
NAANDADLCNLPPKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLLPSRQKK
ADEPDDISTVV

Signal sequence:

amino acids 1-33

Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCAGGTGAGCAAGAGGATGCTGGCGGGG
GGCGTGAGGAGCATGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT
GGGCTCAGTGCTGTGCTCAGGCTCGGCCACGGGCTGCCCCGCTGCGAGTGCTCCGCCCAGG
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG
ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCCGCCAG
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGAGCAACCGCCTGAAGCTCATC
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG
ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT
CATCGTCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC
TGTACCGACTCAAGGTCTTGAGATCTCCACTGGCCCTACTTGGACACCATGACACCCAAC
TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGTTTCTCAACCTCTCCTACAACCCCATCA
GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG
AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTGCGCTCCTGTGGGTGTTCCGG
CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCCGAGTTTGTCCA
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG
CCCGCATCCGGGACCGCAAGGCCCGAGCAGGTGTTTGTGGACGAGGGCCACACGGTGACGTTT
GTGTGCCGGGCCGATGGCGACCCGCGCCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT
ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC
TCCATGCCCCGCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA
GACCTTCGCTTTCATCTCCAACAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG
TGCCTTTCCTTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC
CTGGGCGTCTCCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC
AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCGGGGCGGGGGGAGGGACCCCCG
GGCGGCCGGGAGGGGAAGGGCCTGGTGCACCTGCTCACTCTCCAGTCCTTCCCACCTC
CTCCCTACCCTTCTACACACGTTCTCTTCTCCCTCCCGCTCCGTCCCCTGCTGCCCCCG
CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA
CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA
ATAATTCAATAAAAAAGTTACGAACCTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT
TATGAAAACCTTGAAATAATAAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/199666>; this version posted September 14, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 30

MQVSKRMLAGGVRSMPSPLLACWQPILLVLSVLSGSATGCPPRCECSAQDRAVLCHRKCF
VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFLNLRTL
GLRSNRLKLIPLGVFTGLSNLTQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA
FSGLNSLEQLTLEKCNLTSLIPTEALSHLHGLIVLRRLHNLNINAIIRDYSFKRLYRLKVLEISH
WPYLDTMTFNCCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTLEESVFHSGNLETLILDSNPLA
CDCRLLWVFRRRWRLNFNRRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV
FVDEGHTVQFVCRADGDPPPAILWLSRKHLSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL
CIAANAGGNDSPAHLHVRSSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 556-578

N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,
492-496, 505-509, 526-530, 542-546

Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,
607-611

Tyrosine kinase phosphorylation site.

amino acids 590-598

N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC
ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCCGGGATTTCAGGCTCGCCAGCGCCCAGCC
AGGGAGCCGGCCGGGAAGCGCGAATGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC
TGTTTCGCTGCTGCTGGGCGCCCGGCGGGGCCAACCTCTCCAGGACGACAGCCAGCCCTGG
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA
GAGCCCTTCGAGATAATCGAATTTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTT
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG
CCTGCAGCCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT
ACAGGAAGATCCCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT
GCCCTGATCTTCCCTTTCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG
CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCAGTCCGGTGCCCT
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCTGCTG
CTCATCATGCTCATCTTCTTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTTCATCTAGAGGCGCCTGCCCACTTCTGC
GCCCCCAGGGGGCCCTGTGGGGACTGCTGGGGCCGTACCAACCCGGACTTGTACAGAGCAA
CCGCAGGGCCGCCCCCTCCCGCTTGCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGC
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGAGGGGAGGG
TTGCCCTCAGCCCTTTCCTGGCTTCTCTGCATTTGGGTTATTATTATTTTTGTAACAATCC
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAACA
AACAAAAACA

03009041.074807

FIGURE 32

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDET VVAGGTVVLKCQVKDHEDSSLQW
SNPAQQTLYFG EKRALRDNRIQLVTSTPHEL S I S I S NVALADEGEYTC S IFTMPVRTAKSLV
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRI QEDPNGK
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQR IEVLYTPTAMIRPDPPHPREGQ
KLL LHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMG SYKA
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLT HEAKGSDD
APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,
306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCTCTGG
CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG
TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT
GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA
TCTGTTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA
AGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTACCATTTATTTCTGCATGGCAATTC
CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTAGTTTGCACATGG
AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG
CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA
TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG
ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCCTACCTGCCAAC
GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC
CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT
GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAAACATTTCCCAAGAATGCC
CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC
CACCGAACAGGACTTGTGTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG
CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG
GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT
CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA
GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCTGGGTTTAAAGATGAAC
TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA
GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA
ACCTCATCTGTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG
AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA
GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA
TCCTCCCGGGCACCTTCAATGCCATGCCCAAACCTGAGGATCCTCATTCTCAACAACAACCTG
CTGAGGTCCCTGCCTGTGGACGTGTTCTGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA
CAATTAATCTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA
TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA
GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT
TAGAAAGGATTTTATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT
CGCCACGTTAACTTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC
AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT
GTTTGTACCTCCGCCTTACCCTGGTGGGCATGCTCGTGTATCCTGAGGAACCGAAAGC
GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT
GACTCTTCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCACAGAGTGTATGACTG
TGGCTCTCACTCGCTCTCAGACTAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG
ATACATCCTTCCCCACCGCAGGCACCCCGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG
CCATCAGCCTGGATGGGCATAAGTAGATAAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT
GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA
GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG
AACTGACAGTGCCTTCGCCCTCGCCCCCGGGGCTGTGGGGTTGGATGCCGCGGTTCTATAC
ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCTGTGGATTAG
CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT
TGTAATAAGTAACTTTGACTTCTGAC

FIGURE 34

MLLWILLLETSLCFAAGNVTGDVCKEIKSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH
LFLHGNSLTRLPNEFANFYNAVSLHMENGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKO
TFLGLDDLEYLQADFNLLRDIDPGAQDLNKLVLILNDNLISTLPANVFQYVPITHLDLRG
NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK
IPGNWQIKIRPTAAIATGSSRNKPLANSRPCPGGCSCDHIPGSGMKMNCNNRNVSLLADLKP
KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY
LDTLSREKFAGLQNLLEYLNVEYNAIQILIPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGPNWECSTIVPFKQWAERLGSEVLMSDLKC
ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL
VPGLLLVFVTSFTVVGMLVFILNRKRKRDRDANSSASEINSLQTVCDSSYWHNGPYNADG
AHRVYDCGSHSLSD

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 618-638

N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,
608-612, 657-661, 666-670, 693-697

N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,
354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCTTGACCCCAGAATAACTCAGGGC
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAAACTGTTGGC
CGCTGGGCCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG
AAGGGAGGGGGAACCGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC
AGCTCTGCGTCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG
AGAATGAGGCCGCGTTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCCGGCGG
CGGCGAACACCCCCACTGCCGACCGTGCTGGCTGCTCGGCCCTCGGGGGCCTGCTACAGCCTGC
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG
GCCCCGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT
GCACCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCTCCGACCCCGGCGGTCTC
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC
GGTACTCCAGGCCACCGGTGGGGTGGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGGAGGTCTTGTGTCTGCGCCGCGCCCCGGG
GCCGCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG
TCCACCTGGGACCGAGGTGAGTGCCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA
TCGCGGACGAAATCGGCGCTCGCTGGGACAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA
GTGGGGAAGGACAGCCGACCCTTGGGGGGACCGGGGTGCCACCAGGCGCCCGCCGGCCACT
GCAACCAGCCCCGTGCCGACAGAAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC
ACCACTTGTCCCTGAACAAGACAATTGAGTAACATCTATTCTGAGATTCTCGATGGGGAT
CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC
CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTACGAAAGCCCCTCTTCCCAGCCA
AGGAAGGAGTCTATGGGCCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG
CAGAGGGTGCTTGCTGGCGGAGTCCCCCTTTGGCTCTAGTGATGCATAGGGAAACAGGGGA
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGTATGAAACGGGGAACCAAGAGGAACTTAC
TTGTGTAAGTGAACAATTTCTGCAGAAATCCCCCTTCTCTAAATTCCCTTTACTCCACTGAG
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCCGAGAA
TTTGGAGAAGTGATTGAACTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA
TTAAAAAATAATTTCTACCAAATGGAAAGGAAATGTTCTATGTTGTTTCAAGCTAGGAGTAT
ATTGGTTCGAAATCCCAGGGAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

FIGURE 36

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS
TVRAGAE LRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA
TSPVPQRTWP I RVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 189-193, 381-385

Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,
477-483

Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

FIGURE 37

CGGACGCGTGCGGATTGAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC
CCCCGGCTACCTGGGCCGCCCCGCGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGCAGCCGA
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGCGGTGTGAGCGCGGTGGGTGCGGA
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCGAGCGTCTACGCTGCCATGA
GGGGCGCGAACGCCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG
CAGCAGTCCCCAGAGAGACCTGTTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT
TATTGGCAGTGAAGGTTTTCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAATCA
CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATACTACTGC
CGATATGATTATGTGGCTGTGTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTTATTGGTCACTACATATTAGGCCAAAA
AAACTGCCTACAACACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAACCATCACTCGCGATGGGAGTTTG
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTACAGCAGGCGGG
CAAGAACATGAGTGCCAGGCTGACTGTGCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG
TTAAACAGTGAAGTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTT
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCGAAAGATGG
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT
TGCCTGCTGTGAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA
AGCTAAAAGTGTCAAGCGTTGACAGCTTGAAGCGTTTATTTATACATCTCTGTAAAAGGAT
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGT
GTTATTTGTTTACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA
AATCAATGCTTAATAAAATATTTTTTAAAGGAAAAAAAAAAAA

FIGURE 38

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTECGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVLNFRFIDLESNLCRYDFVDVYNHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPAPIVSERNELLI
QFLSDLSTADGFIGHYIFRPKKLPPTTTEQPVTTTTFPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLACTVITTTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,
305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

FIGURE 39

CGGACGCGTGGGCGGACGCGTGGGCGGCCCCACGGCGCCCGCGGGCTGGGGCGGTGCTTCTT
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGGCCCCATGGCCCCGAAGG
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGG
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG
GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTTACAAGCAGCAGGAGGCCCCGGACCTCTTCC
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC
CTTCCCTGTCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTAAGGAGAAGG
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT
TTTGGCCCCCTGTGCCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGG
CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA
ACAAGCAGTGTGAAAAACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC
AGAAGACGAGTTGGTGGTGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA
CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG
ACTGGCTACTGGTTGTGAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCC
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG
CCTTACAGAGCAGCCCAGGTACCCAGGCCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC
CCTGAAGGTGGATAACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA
ATTTCAAAAGTTTTTCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT
TCTGTGTTACACATCCCCACACCCCATTGCCACTTATTTATTCATCTCAGGAAATAAAGA
AAGGTCTTGAAAGTTAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 40

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER
TIRDNFGGGNTAWEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG
YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPPEESNCLQCKKGWALHHLKCVDIDE
CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDLVVLQQMFFG
IIICALATLAAGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

Signal sequence:

amino acids 1-29

Transmembrane domain:

amino acids 372-395

N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,
326-332, 372-378, 395-401

Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

EGF-like domain cysteine pattern signature.

amino acids 181-193

FIGURE 41

5'-TGAGACCCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA
GCACCATGCAGCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGG
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT
GCCACCCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGGCCAGT
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGAAAGAGGTTTCAGCCAGAGC
TTCCGAGAGGTGGCCGGCAGGTTCTTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG
CATGGAGCAGCGGCTGCCGCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCGGGCC
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCCGTGCAGAGG
GAGCATCTGGGCCCCGCTGGCGTCCGGCGCCCAACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG
CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG
ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATAGGCGCCTAGTG
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTGAGGGTACCAGGAGAGCTG
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT
CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTACTGCACCTATATTCTAAGCACTTACAT
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCANTGTGTATTGTTTACTTGTCTGTAC
TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT
AAAA

FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV
ALLQRSHGDRSRGKRFSQSFRVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP
VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCTCRQPPEALAFKWPF
LGPRQCIASETDSLPMIVSIKEGGTRPQVVSLPNMRVQKSCASDGALVPRRLQP

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.11.336660>; this version posted March 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 43

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCA GTGGCCTGATCGCGATGGGGACAAA
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAAC TGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGT CATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTTCCTGGTGTGAGCCTGGTCTGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTGAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTTCTACCACTGCTGAGTGGCCTGGAACCTTGTTTTAA
GTGTTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTCTGCAGGAATCTGCACTCAACTGCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAAC TGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCTGGGAGTTCTGGGATCAGCCTGACCA
ACATGGAGAAACCCTACTGGAAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

FIGURE 44

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI
LLGILVFGIWFAYS RGHFDR TKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

FIGURE 45

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCT
GGAGGCCGCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG
GCTCGTGCCCAACCAAGTTCCAGTGCCGCACCAAGTGGCTTATGCGTGCCCTCACCTGG
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC
ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA
GTGACTGCTCTGGGGGAACTGACAAGAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTACCTCTCTCAGGAATGCC
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTGCGGAATGCCACATCCTC
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCTGGCTCCGAGCCCAGGAGCGC
CTCCGCCCCACTGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC
CTCGCTGCCCTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAAGTGGCCCTGGAGATTGAGGGTCCC
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

FIGURE 46

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR
TSGLCVPLTWRCRDLDCSDGSDEEEECRIEPTQKGQCPPPGLPCPCTGVSDCSGGTDKKL
RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

Signal sequence:

amino acids 1-30

Transmembrane domain:

amino acids 230-246

N-glycosylation site.

amino acids 126-130, 195-199, 213-217

Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,
224-230, 230-236, 263-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

Leucine zipper pattern.

amino acids 17-39

FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGATCATGGAATGAACCCGAGCAATG
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCGAGCG
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCAGTCCAAGACGGATTCAAGCTGAAG
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTATA
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG
TACCCCGACCTACACAATATGGTTTCATTATGTGCGGATGATGGAACGTGGAATAATCTGCC
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAACTT
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCGGTG
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGTCT
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAAGTGTGGTGGAGTTTTACTGCGATCCT
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGTTTCCTTC
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAG
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG
ACGAAGCTGTGAGTGGCGGCTTGAGTGCCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG
GGCTGCCCTTACCCGTGGACGACCAGAGCCCCCAGCATACCCCGGCTCAGGGGACACGGA
CACAGGCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCACTGGGT
GTTGTTCTTAAGAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA
ATACATGTTGATCTGTGGAGTTGATTCTTTCTTCTTGGTTTTAGACAAATGTAAACAA
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAAGNAAAAAA

[illegible]

Signal sequence:
amino acids 1-41

N-glycosylation site.
amino acids 104-108, 134-138, 192-196

N-myristoylation site.
amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,
478-484

Prokaryotic membrane lipoprotein lipid attachment site.
amino acids 394-405

[illegible]

CCCACGCGTCCGCTCCGCGCCCTCCCCCCCCGCTCCCGTGCGGTCCGTCCGGTGGCCCTAGAGA
TGCTGCTGCCGCGGTTGCAGTTGTGCGGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT
AGCGCCCGAGTGTCTGGGGGGCGCACCCGAGTCGGGGCCATGAGGCCGGGAACCGCGCTACAGG
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC
TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAA
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT
CACAAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTTCATG
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTTACATGTTCCAGTGGAATGATGA
CCGGTGCAACATGAAGAACAATTTCAATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG
GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT
CCTAATCCCCAGCATTCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCAATCCGAGTGTGTT
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG
GTTATTAGGACATATAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAATC
CTCTTATTTTCTATAAGGAAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTTGGCTGTATCCTTTAT
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA
GAGTCTCAATAAATGTCACCTTGGTTGGTTGTATCTAACTTTTAAAGGGACAGAGCTTTACCTG
GCAGTGATAAAGATGGGCTGTGGAGCTTGGAACCACCTCTGTTTTCTTGCTCTATACAG
CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCACATATGGTAGCACAG
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA
GCAGGAAAAA

FIGURE 50

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLLRGGQPVCRRGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGD FWIGLRRREEKQSNSTACQDL
YAWTDGSISQFRN WYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLLV
VTTVVCWWICRKRKREQPD PSTKKQHTI WPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPD DMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

Signal sequence:

amino acids 1-21

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,
299-303, 306-310, 323-327

N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT
GTTCAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAATC
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT
GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC
TTTACCAACCTTCAACCCACAGGAGGTCTTTATTTCGTTCCACTAACATTTTTTCGGAATCTG
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT
CCCACTGATGAAGCAGATTTCAGAAGTCTTGATCCCAACTACCAAAGCTGCTGGAGCCTGA
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT
GTTGACCTGACCATGGAACCTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCCGCTGGACATGT
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA
ACTCAGGTGATGGAAGTTGGAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT
TTAAAATAAAGTGCCTTTATACAATG

FIGURE 52

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMOVVFRHGAR
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL
TKVGMQOMFALGERLRKKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFCQKE
GPIIIHTDEADSEVLYPNYQSCWSLRQRTGRRTASLQPGISEDLKKVKDRMGIDSSDKVD
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFHILES
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLMTLGIFDHKWPPFAVDLTMELYQHLESKEW
FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMS VYTLSPEKYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 218-222

Casein kinase II phosphorylation site.

amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site.

amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature.

amino acids 50-65

FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT
TAAATTTTCAGCTCATCACCTTCACCTGCCTTGGTCAATGGCTCTGCTATTCTCCTTGATCCTT
GCCATTTGCACCAGACCTGGATTCCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCTTCATCCA
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA
GTCCAGAGGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCGCGCAAAGGTGG
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA
GGATTGCCCTTCTGGGCCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTGCG
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTTCGTTGCTCAGGGGAGGAG
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA
TGTGGCTGTCATCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCTTCAGAGTTGG
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACCTACATCA
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA
CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC
TATGTGTTCAAA

FIGURE 54

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC
RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCCQTGWSLRAAKVVCRQLGCGRAVL
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG
GDNLCSGRLEVLHKGWVGSVCDNNGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

Signal sequence:

amino acids 1-15

Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,
267-271, 294-298, 316-320, 336-340

N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,
180-186, 263-269, 286-292

Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG
GTGAAGGCCATGGACTTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCTGGCTGCCTGGG
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC
GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA
TGTGGACAAGAGGGTTCATGGAGACAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG
ATGAGCATTCCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGCTGTTTATCTTCGAA
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA
AGCTTCTTCCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT
AAAACTAGAAATAAACATCTCAAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

[illegible]

Signal sequence:

Transmembrane domain:

N-glycosylation site.

Glycosaminoglycan attachment site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA
AAAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTTAA
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA
GGCATTCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG
CTGGACATGTCTCGGTCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA
TTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG
TCTGTGTCCTAATTTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG
ATTTTTATTCCATCTTCTATAGCTTTTTTAAACAACATTGGAAAGGATCCTTCCTGAGCGTTT
CCTGGCAGTTTTTAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA
AAGCGCAATAAGCACCTAGTTTTCTGAAAACCTGATTTACCAGGTTTAGGTTGATGTCATCTA
ATAGTGCCAGAATTTTAATGTTTGAACCTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA
CCAAAATGACTTTATTAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTGT
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA
TGAAGGACTATATCTAGTGGTATTTTCAATGAATATCATGAACTCTCAATGGGTAGGTTTC
ATCCTACCCATTGCCACTCTGTTTCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA
AA
AA

0909054 074304

[illegible]

Signal sequence:

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

[illegible]

CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCCGCCGCGGCTC
AGGGAGGAGCACCGACTGCGCCGCACCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG
TTTCGCTGGTCCCTGTTGATGCCTGGCCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC
TGGGAAGATCCAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA
AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGTTCT
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGCTACAGGGTGGGCCGGG
AGGTTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA
CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT
AGCACGGGATTTATACAGTGCCTAATTGAGTTTTTCCAGATATTTCTGAATATAAAAAATA
ATGACTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCTGTACCAAATTGGCT
TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC
AGGAAGCAGAAGCTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAA
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTTGCGGT
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCCTCCAGAGGTGAGACAA
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACATAGTTGAAAAGTACTTGCGAGA
AGATACAGTAGCACTAGTTAAGCCATTGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA
TCTACAATGCCCCACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC
ATGGACTGGAAGGATCCAGGAATACAAGAAGGCAGAAAAAAAAGTTTGAAGATCTTTAA
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA
TTCATTTATGGAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCAAAAGAGAACAT
CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA
TTATCTTTTCATATCTGCAAGATTTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC
TTTTGTTTTTGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA
AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT
GTTTGGAAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA
TAGTTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA
AAAAATATTATATATAAAAGTAAAAAAAAAAAA

1. *Staphylococcus aureus* (ATCC 12228) was grown in tryptic soy broth (TSB) (Difco) supplemented with 0.5% yeast extract (Difco) and 0.5% glucose (Difco) at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with phosphate-buffered saline (PBS) (pH 7.4). The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier (Branson Ultrasonics, Danbury, CT) to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

2. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

3. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

4. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

5. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

6. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

7. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

8. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

9. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

10. *Staphylococcus aureus* (ATCC 12228) was grown in TSB supplemented with 0.5% yeast extract and 0.5% glucose at 37°C. The cells were harvested at mid-log phase (OD₆₀₀ = 0.5) and washed with PBS. The cells were then resuspended in PBS and sonicated using a Branson 250 sonifier to release the cell wall components. The cell wall components were then purified by dialysis into distilled water and lyophilized.

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTGGCTACAACAT
TTTTCCCTTTTCTAACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC
CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG
TGAACCTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA
TAGAGATGCTTTGTAAAAATAAAATTTTAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA
GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGAGAAAGTATGTTAAAAATA
GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTGGTCTGACCACTCTGCCTTGTGTTTT
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT
CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG
ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCCAGCAGCCGGC
ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCCTTGACCGT
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA
ACCTGACCATCCAGGTGGCTCATAAGACAGGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG
CCCCTCATCGTGAGCCCTGCAGCGAAGTGCTCACCTCACCAACAATGTCAACAAGCTGCT
CATCATTGACTACTCTGAGAACCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA
AGCTGCTGCGGCTGGATGACCTCTTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC
CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA
GCCGGAAGCTGCCCCGAGACCCTGAGTCTCAGCCATGCTCGACTATGAGCTACACAGCGAT
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT
CTTCTACATCTACGGCTTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG
CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC
AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG
CAGTATCACCAACCGCCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT
GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC
TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC
ATTACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG
GCAACTTTATTTTCTTGGGGAACAAAGGTGAAATGGGGAGGTAAGAAGGGGTTAATTTTGTG
ACTTAGCTTCTAGCTACTTCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA
TTTCAATATTTCCCAAACCTTTAAGAAAAAATTTAAGAAGGTACATCTGCAAAAGCAAA

FIGURE 62

MGTLGQASLFAPPGNYFWSHDHSAFCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPLIVQPCSEVLTLTNVNKLLIIDYSENRLACGSL
YQGVCKLLRLDDLFI LVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPDPPESSAMLDYELHSDVSSLIKIPSDTLALVSHFDIFYIYGFASSGGFVYFL
TVQPETPEGVAINSAAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSAFCFAPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLLKVRVYEFRCSSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

Signal sequence:

amino acids 1-32

Transmembrane domain:

amino acids 71-87

N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,
384-388, 471-475, 481-485, 530-534

N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGATGACCA
CGTGAGCCTCCGGCGGAGGCCCGGCCGACGCTGGGACTCCTGCTGCTGGTCTGTTGGGCTTCCTGGTGCTCC
GCAGGCTGGACTGGAGCACCTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTCGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT
GGAGGGACCGCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCCTCACCACCTATGTTCCGTGGAACCTGCATG
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCCTTCGTCTGATGGCCGCAGAGATCG
GGCTGTGGGTGATTCTGCGTCCAGGCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCAGCTGGCTAC
TCCAAGACCCTGGCATGAGGCTGAGGACAACCTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG
GTTCTTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACATGC
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGTGAGT
ACTGACCGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG
GAGCCATGCATTTCCATGACTACAAGTCAGATGTCAACAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACTGCCAGTCAATGGGGGAAATGGACAGTCCCT
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT
ACACCGTGTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG
GCTTAATTGGAAATCTCTATCTGAATGATTCAACCCCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA
GCTTCTTTTCAAGGTTTCGGCCTGGACAAATGGNGTTCCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTGG
GTAGCTTGTCCATCAGCTCCACGCCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATTCA
TCAATGGCCAGAACCCTTGGACGTTACTGGAACATTGGACCCCAAGACGCTTTACCTCCCAGGTCCCTGGTTGA
GCAGCGGAATCAACCAGGTCATCGTTTTTGGAGAGACGATGGCGGGCCCTGCATTACAGTTACAGGAAACCCCC
ACCTGGGCAGGAACCACTACATTAAGTGGAGCGGTGGCACCCCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG
ACTGGGGGCTACAGTCTGCCCCCTGTCTCAGCTCAAAACCCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC
TGGCTTTGTTGATGATGGCTTTCCCTACAGCCCTGCTCTTGTGCGGAGGCTGTGCGGCTGTCTCTAGGGTGGGAGC
AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGCACCGGACGTACAGCCC
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTGTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT
TTTATCCCCGAAATCCTGGGTGTGTCAACAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT
CTTCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACTCGGCGTGAGAAACATGTGACTTCCCCTT
TCCCTTCCCACTCGCTGCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC
CAAGTTAGCAGGTGTCTCTGTTGTTCAAGTGGAGGACATGTGAGTCTTGGCAGAAGCCATGGCCCCATGTCTGCA
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCCATGTCTGCACATCC
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCCATGTCTGCACATCCAGGGA
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCCATGTCTGCACATCCAGGGAGGAGG
ACAGAAGGCCAGCTCAGTGGCCCCGCTCCCCACCCCCACGCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC
GAAGTGTGTCCAAGTCCGATTTGAGCCTTGTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCACTGTCTCTGA
GTTGCAGTAAAGCTATAACCTTGAATCACAA

THE FEDERAL BUREAU OF INVESTIGATION

Signal sequence:

Casein kinase II phosphorylation site.

N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

Abstract—The purpose of this study was to determine the effect of a 10-week training program on the heart rate (HR) and energy expenditure (EE) of sedentary, middle-aged men. The subjects were 10 men, 40 to 50 years of age, who were sedentary and had no cardiovascular disease. They were randomly assigned to a 10-week training program or a control group. The training program consisted of 30 minutes of aerobic exercise, 3 times per week, at a heart rate of 150 beats per minute. The control group did not exercise. The HR and EE were measured at rest and during exercise at the beginning and end of the 10-week period. The HR and EE were significantly higher in the training group than in the control group at the end of the 10-week period. The results of this study suggest that a 10-week training program can improve the HR and EE of sedentary, middle-aged men.

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGCACCTACGCCAAAGG
GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCC
AAGAAGCTGTCTTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCA
GGCAGACACTCGGTGCTTCGTAGTGATAGGGGT CATGACCGGTTTCTCTAGACGGGGCCC
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAAC TA
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATT CATCTAAGAACCTC
AGATCCAGACTTCTTTGCCGCAGTGGACTCCTGGTTC AAGGTCTTGCTGCCCAAGATATATC
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATT CAGGTGGAGAATGAATATGGTAGC
TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCGCTGCACTGCTAGG
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA
CTGGGGCCAGAATCACTCCACACGGTCTGTGTG CAGCTGTAACCAAGGACTAGAGAACATGC
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG
AATGGTGCCGATAAGAAGGGACGCTTCCTTCG GATTACTACCAGCTATGACTATGATGCACC
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTT TGTCTCTTCGAGATGTCATCAGCAAGT
TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCG GAGCCCCAAGATGATGCTTGGACCTGTG
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTT CCTAGACTTGCTTTGCCCCCGTGGGCCCAT
TCATTCAATCTTGCCAAATGACCTTTTGAGGCTGT CAAGCAGGACCATGGCTTCATGTTGTACC
GAACCTATATGACCCATACCATTTT TGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC
CATGACCGTGCCTATGTGATGTTGGATGGGGTGTT CCAGGGTGTTGTGGAGCGAAATATGAG
AGACAACTATTTT TGACGGGGAAACTGGGGTCCAA ACTGGATATCTTGGTGGAGAACATGG
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTT CAAGGGCCTGTTGAAGCCACCAATTCTG
GGGCAAACAATTCCTTACCCAGTGGATGATGTTCC CTCTGAAAATTGATAACCTTGTGAAGTG
GTGGTTTCCCTCCAGTTGCCAAAATGGCCATATCCT CAAGCTCCTTCTGGCCCCACATTCT
ACTCCAAAACATTTCCAATTTT TAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTT TGGGCCGGTACTGGACAAAGCAGGGGCC
ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTT TCCTAGGGGAGCCCTCAACAAAATTA
CATTGCTGGAAC TAGAAGATGTACCTCTCCAGCCCCA AGTCCAATTTTTGGATAAGCCTATC
CTCAATAGCACTAGTACTTTGCACAGGACACATAT CAATTCCCTTT CAGCTGATACACTGAG
TGCTCTGAACCAATGGAGTTAAGTGGGCAC TGAAGGTAGGCCGGGCATGGTGGCTCATGC
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGG ATTACCTGAGGT CAGGACTTCAAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTA AAAAATACAAAATTAGCCGGGCGTG
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCT GAGGGCAGGAGAATTGCTTGAATCC
AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACC ACTGCACTCCAGCCTGGCTGACAGTGA
GACACTCCATCTCAAAAAAAAAAAAAA

[illegible]

Signal sequence:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC
ACCCACAATATGGCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTTGGTGTGTT
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG
AAAACTCAGGCAGCACATTTACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTTCATG
CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA
TGCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTCATAATGACGGCAC
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTGCTGAGCTGGAACCTCCAGA
ACTGTGAGCTAGAGAGAATCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACCTGGAT
TTAAAGTCCAATAACATTTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTATAAACGACT
GACTTGTTTAAAAATTATGGCATAACAAAATTGTTACTATTCCCTCCCTCTATTACCCATGTCA
AAAACCTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA
AATAGGATTGCTTCAGAACCTGCAGCATTTCATATCACTGGGAACAAAGTGGACATTCTGC
CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAAC
CTTGACCGCCTGCCAGCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG
TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA
AATATTCCCTTTGCAAATGGGATTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTCTTTTTTAAATTGTTTGTAACTTGGAT
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT
TTCTTACTAAAAA

2025

Signal sequence:

N-glycosylation site.

cAMP- and cGMP-dependent protein kinase phosphorylation site.

Casein kinase II phosphorylation site.

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

FIGURE 69

CCCACGCGTCCGGCCTTCTCTCTGGACTTTGCATTTCCATTCCCTTTTCATTGACAACTGACTTTTTTTTATTTCT
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT
CTGTGTTTTGGGGTTTCTTCTCCCTCCCTGACATTGGCATTGCTTAGTGGTTGTGTGGGGAGGGAGACCACGTGG
GCTCAGTGCCTTGCCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC
ATCGCTGGTGGTATCCTGGCGGCCCTTGCTCCTGCTGATAGTTGTGCTGCTCTGTCTTTACTTCAAAATACACAAC
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCAGACAAGGTGTGGTGGGCCAAG
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCTGCTGCTGCAAGTGTGAAGGATATAGAATGTGTGCC
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAA
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG
CTCCCCATCAGTTTTCATGGAAAAATACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC
CACTGGGGGCAACCCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCCTCACTGTGGGGAATGCTGATAAACCAAGTCA
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAAGTACGTTTTCCCTGGAGGTGTCCAGAAA
GCTGATGTAACACAGAGCCTATAAAAGCTGTGCGTCCCTAAGGCTGCCAGCGCCTTGCCAAAATGGAGCTTGTA
AGAAGGCTCATGCCATTGACCCCTTAAATCTCTCTCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT
GCAAGCTGCACAGTCACTTAGGGGGTGCCAAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC
AGTGAGAACTGCACCTGGACAATAGAAAAGACCAGAAAAACAAAGCATCAGAATTATCTTTTCTATGTCCAGCTT
GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCCTCTGCTAGGG
CAAGTCTGCAGTAAAAACGACTATGTTCCCTGTATTTGAAATCATCATCCAGTACATTGACGTTTTCAAATAGTTACT
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCTAATCATCTCTATTCCAACTGT
GGCGGTTACCTGGATACCTTGAAGGATCCTTACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCTAGAAATAGAC
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT
GGCCGTGTGACTCCACCTTCGAATCGTCATCAAACCTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT
TACCGGGGATTTTCTGCTTCCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA
GACCCAACTTGACAGACCAAAATTTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC
CGTCAGAAACAACCTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA
GAAGATGATGTAATACAAAGTCAAAATGCCTGGGCAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA
TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTGTAACCAAACCTCTTTTGTTCAGTTAGTCTGCAC
ACCTCAGATCCAAATTTGGTGGTGTCTTGTGATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC
TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAAGTGTATCCCTTATTTGGACACTATGGGAGA
TTCCAGTTTAAATGCCTTTAAATTCCTGAGAAGTATGAGCTCTGTGATCTGCAGTGTAAGTTTTGATATGTGAT
AGCAGTGACCACCAGTCTCGCTGCAATCAAGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG
AAAACAGATTCATCATAGGACCCATTCTGCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTCAGCAT
GAAACACATGCGGAAGAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCTTTCATGGTTCTAGCTCTG
AATGTGGTGAAGTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG
CAGAACTATTAACTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT
GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

FIGURE 70

MELVRRRLMPLTLLILSCLAELTMAEAEAGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI
ERPENKSIRIIFSIVQLDPDGSCSENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSTLT
FQIVTDSARIQRTVFVFYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV
EKDYKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVL
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIIISKSYLEAFNSNGNNLQLKDPTCRP
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSSIIGPIRLKRDRSASGNSGFQHETHA
EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

Signal sequence:

amino acids 1-24

Transmembrane domain:

amino acids 571-586

N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,
394-398, 419-423

Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,
408-412, 463-467, 520-524, 556-560

Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

FIGURE 71

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGC
TGCTGCCGCCGCCGCGTGCCTGCCACAGCGCCACGCGCTTCGACCCACCTGGGAGTCC
CTGGACGCCCGCCAGCTGCCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG
GGGAGTGTTCCTGCGTGCCTCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCCTAGTTTCAAATATGAAGAT
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTTCAGGC
CTCTGGTGCCAAATACATTGTCTTAACCTTCAAACATCATGAAGGCTTTACCTTGTGGGGGT
CAGAATATTCGTGGAACCTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA
CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTTGACTGTACTATTCCTTTTTTGA
ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACATATCAGCCTGAGGTTCTGTGG
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT
ATATAATGAAAGCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA
CATAAATGGGAAAACCTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAAGCTGG
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG
GAGGAAATCTTTTGATGAATATTGGGGCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG
GAGCGACTGAGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA
TACCTGGCGATCCCAGAATGACACTGTCACCCACAGATGTGTGGTACACATCCAAGCCTAAAG
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCACATCAGGACAGCTGTTCCCTTGCCAT
CCCAAAGCTATTTCTGGGGGCAACAGAGGTGAACTACTGGGCCATGGACAGCCACTTAACTG
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTTCATCAGATGC
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTG
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA
AAGCAATGTAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTA
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACCTTTGCCATTAAAGTC
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTTATGTTGAAGCCATATCCCCCATG
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTTCTGGGAAATGCATTGCTAGTCAAT
TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACCTTTACAAACGTTTTAATT
TTTAAACCTTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAACATAAACTCATTGTGCAA
ATGTAA

FIGURE 72

MRPQELPRLAFPLLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEFWFWYWQKEKIPKYVEFMKDNYPSPFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKKGWALALTNVI

Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,
375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT
TTCCATCCAGGTGTCTATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGC
AGCTGAATTTACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGGAGA
TGGATTTCGTGGTCACTCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG
TCCTGATTTTGAAGGTTCCAGTGAGCCGACAGTTTTGCAGCCTATTGTTACAACCTCATCTGAT
ACTTGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA
AACTGCAACACAAACAACAGAATTTATTGTCTAGTGACAGTACCTACTCGGTGGCATCCCCTT
ACTCTACAATACCTGCCCCCTACTACTCTCCTGCTCCAGCTTCCACTTCTATTCCACGG
AGAAAAAAATTGATTTGTGTACAGAAGTTTTTATGGAACTAGCACCATGTCTACAGAAAC
TGAACCATTGTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC
CCACGGCTCTGCTAGTGCTTGTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGGATTTTGC
TATGTCAAAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGCATGCCTGGAA
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTTTCATGCTCC
TTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAAAGAAGAAAGTCCACCCTT
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC
CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC
TTTCTAGCCTGGCTATGTCTAATAATATCCCACTGGGAGAAAGGAGTTTTTGCAAAGTGCAA
GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG
TGGGTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC
CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA
AGAGCAAAAAGTGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT
AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTTTATCTGA
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT
TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTTCA
GTTTTGATATTTCTAGCTTATCTACTTCCAACTAATTTTTTATTTTGTGCTGAGACTAATCTT
ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACTAAGAAG
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA
GCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA
AGCATTTAGAAAACCTT

FIGURE 74

MARCFSLVLLLLTSIWTTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR
LLGLSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLWKVPVSRQF
AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP
APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNAAAGFGGVPTALLVLALLFF
GAAAGLGFCYVKRYVKAFPFITNKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP
SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 75

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTTCGGTGCCGCGACTTTCACGATGG
CTCGCCCAACCTTACTACCTTCTGTCTGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA
ACTGCCGCCGCTCTGCCACGGTCTGCCACCCCAACGCGAAGACGGTAACCCGTGTGACTTTG
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT
TCTTTTCTTCCGCTTGGATATTTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC
TGATGACGTGCAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA
TTGGTCTAATGACTGCCAATCATTGCCCCCTATCTATGCTGACCTCTCCCTTAAATACAAC
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG
AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTCAGA
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCCTCTCCTGTCAATT
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGCAAGAAAGAGATC
TCATAGGACGGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTGAGTGTGTTAACTGCTTATC
AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCTTCTTAG
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC
TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCGTAATAAAAAGATTGGGATT
TCCTTTTG

FIGURE 76

MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLLRKLPPLCHGLPTQREDGNPCDFD
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

Signal sequence:

amino acids 1-48

Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

FIGURE 77

GGACAGCTCGCGGCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG
GGGCCCCAGCCTGGCCCCGGGTCAACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA
TTGCTCCTGCTGCCCCGGCTCCTACGGA CTGCCCTTCTACAACGGCTTCTACTACTCCAACAG
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG
TGGTGGAGACACCCGAGGAGACCCTGTTACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAAGCT
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCA TTGACGGGCTGGA
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA
ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCGGCAGCCCTGCGGTGGCC
CAGGCCTGGCACCTGGCGTGC GAAGCTACGGCCCCCGCCACCGCCGCTGCACCGCTATGAT
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCTGAGAAGCT
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC
AGCTCTTTGCCGCCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT
GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGT TTTACTGCTACCGCCAGC
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT
CGTTTTCCCTTGTGGGTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT
TTAAACATTTTTTTTACTATTTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTTGTGGGCTTCTG
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGGCCAGAGTGGGC
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCCT
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC
CACCTCTACTTCTCTGTGAAGCCGCTGACCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAACTCCCCTCCCCGTCCCCCT
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTGTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA
AAA
AAA

FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD
KEHDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR
HRR LHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR
CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC
CAGGGGAGGGTGCACCAGGCGGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA
CTTCCAGTACGACCATGAGGCTTTCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA
CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTTCATGAC
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCTGCACCCCGAGG
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC
CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA
AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG
ACCTGACCCGGCACCCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG
CACAATGACCGGAGGAGGGGCCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCGCAGGAG
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCCTGGGCTCTCAGGGACCCCCTGGGTGCGGCTTC
TGTCCCTGTACACCCCCCAACCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC
CTATTTCTGACTGAGTCTCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA
AAGCCAGCGCCGGGACCTTGAAA
AAAAAAAAAAAAAAAAA

FIGURE 80

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR
EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT
YDTRDRGVRVGWEELRNATYGHYAPGEEFHDVEDAETYYKMLARDERRFRVADQDGDSMATRE
ELTAFLHPPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAAWVQTERQQ
FRDFRDLNKDGHLDGSEVGHVWLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV
GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,
291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

[illegible]

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG
GCGGCGGGGCGCGGGTGCAGGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCG
CCTGTCTGTTCGTTCGTTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTGACACACCGATCCTG
GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG
CGGGTCGTCTGTGTCTCTCTCCTCTGCGCCGCGCCCCGGGGATCCGAAGGGTGCGGGGCTCT
GAGGAGGTGACGCGCGGGGCCTCCCGCACCCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCG
GTGTGAGCAGCCTATCAGTCACCAATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG
TGTCTGCTGCTGCTGCCGGGGGCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG
TTTTACCAGAGGCTTGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC
TGGTCGAGAAAACCTATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT
GGTCTGCTTCTTTTACAGTAACTAAAGGC AAAAGTAGTACACAGGAGGCCACAGGACAAGCA
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGG
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC
GCCGATTTAATTTACAGAAGAATTTTGTGTGAAAAGTGGCTCTAATGTTGGGAATTGGAACA
GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAATAGAATTTTACTTGAA
AACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTATTATTGATGGTTGGCCTTCTGATGACAT
CGAGGAAGCAGGCATTGTG GCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA
AGCCTATCCCTGAAGAACTGGGGATGGTTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGT
CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTGAGGATAGCAATTTCCGCCTC
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA
AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT
GATGCCATTTCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAAAGAA
CTTCTTAGTAATTGTACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG
ACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCCTTGGGCACCTCTGGATGACCTG
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTCTTCAACAAGAGAGTTTCACAGGATT
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT
AATG GTAACATTTTGACAACCTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT
CTCATAATACTGAAATGCTTTAGCATATAAGATACAGATACAAAACTATTAAAGTATGTCAAC
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT
ACTTTGTTAAAAACACTGCTGAGGCTTATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA
GATAATGTGGATTAAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA
TTCCATAGCTCAATAAAAAGAATCTGATACTTAGACCAAAAAAAAAAAAA

1. **Introduction**
 2. **Background**
 3. **Methods**
 4. **Results**
 5. **Conclusion**
 6. **References**
 7. **Appendix**
 8. **Figure 1**
 9. **Figure 2**
 10. **Figure 3**
 11. **Figure 4**
 12. **Figure 5**
 13. **Figure 6**
 14. **Figure 7**
 15. **Figure 8**
 16. **Figure 9**
 17. **Figure 10**
 18. **Figure 11**
 19. **Figure 12**
 20. **Figure 13**
 21. **Figure 14**
 22. **Figure 15**
 23. **Figure 16**
 24. **Figure 17**
 25. **Figure 18**
 26. **Figure 19**
 27. **Figure 20**
 28. **Figure 21**
 29. **Figure 22**
 30. **Figure 23**
 31. **Figure 24**
 32. **Figure 25**
 33. **Figure 26**
 34. **Figure 27**
 35. **Figure 28**
 36. **Figure 29**
 37. **Figure 30**
 38. **Figure 31**
 39. **Figure 32**
 40. **Figure 33**
 41. **Figure 34**
 42. **Figure 35**
 43. **Figure 36**
 44. **Figure 37**
 45. **Figure 38**
 46. **Figure 39**
 47. **Figure 40**
 48. **Figure 41**
 49. **Figure 42**
 50. **Figure 43**
 51. **Figure 44**
 52. **Figure 45**
 53. **Figure 46**
 54. **Figure 47**
 55. **Figure 48**
 56. **Figure 49**
 57. **Figure 50**
 58. **Figure 51**
 59. **Figure 52**
 60. **Figure 53**
 61. **Figure 54**
 62. **Figure 55**
 63. **Figure 56**
 64. **Figure 57**
 65. **Figure 58**
 66. **Figure 59**
 67. **Figure 60**
 68. **Figure 61**
 69. **Figure 62**
 70. **Figure 63**
 71. **Figure 64**
 72. **Figure 65**
 73. **Figure 66**
 74. **Figure 67**
 75. **Figure 68**
 76. **Figure 69**
 77. **Figure 70**
 78. **Figure 71**
 79. **Figure 72**
 80. **Figure 73**
 81. **Figure 74**
 82. **Figure 75**
 83. **Figure 76**
 84. **Figure 77**
 85. **Figure 78**
 86. **Figure 79**
 87. **Figure 80**
 88. **Figure 81**
 89. **Figure 82**
 90. **Figure 83**
 91. **Figure 84**
 92. **Figure 85**
 93. **Figure 86**
 94. **Figure 87**
 95. **Figure 88**
 96. **Figure 89**
 97. **Figure 90**
 98. **Figure 91**
 99. **Figure 92**
 100. **Figure 93**
 101. **Figure 94**
 102. **Figure 95**
 103. **Figure 96**
 104. **Figure 97**
 105. **Figure 98**
 106. **Figure 99**
 107. **Figure 100**
 108. **Figure 101**
 109. **Figure 102**
 110. **Figure 103**
 111. **Figure 104**
 112. **Figure 105**
 113. **Figure 106**
 114. **Figure 107**
 115. **Figure 108**
 116. **Figure 109**
 117. **Figure 110**
 118. **Figure 111**
 119. **Figure 112**
 120. **Figure 113**
 121. **Figure 114**
 122. **Figure 115**
 123. **Figure 116**
 124. **Figure 117**
 125. **Figure 118**
 126. **Figure 119**
 127. **Figure 120**
 128. **Figure 121**
 129. **Figure 122**
 130. **Figure 123**
 131. **Figure 124**
 132. **Figure 125**
 133. **Figure 126**
 134. **Figure 127**
 135. **Figure 128**
 136. **Figure 129**
 137. **Figure 130**
 138. **Figure 131**
 139. **Figure 132**
 140. **Figure 133**
 141. **Figure 134**
 142. **Figure 135**
 143. **Figure 136**
 144. **Figure 137**
 145. **Figure 138**
 146. **Figure 139**
 147. **Figure 140**
 148. **Figure 141**
 149. **Figure 142**
 150. **Figure 143**
 151. **Figure 144**
 152. **Figure 145**
 153. **Figure 146**
 154. **Figure 147**
 155. **Figure 148**
 156. **Figure 149**
 157. **Figure 150**
 158. **Figure 151**
 159. **Figure 152**
 160. **Figure 153**
 161. **Figure 154**
 162. **Figure 155**
 163. **Figure 156**
 164. **Figure 157**
 165. **Figure 158**
 166. **Figure 159**
 167. **Figure 160**
 168. **Figure 161**
 169. **Figure 162**
 170. **Figure 163**
 171. **Figure 164**
 172. **Figure 165**
 173. **Figure 166**
 174. **Figure 167**
 175. **Figure 168**
 176. **Figure 169**
 177. **Figure 170**
 178. **Figure 171**
 179. **Figure 172**
 180. **Figure 173**
 181. **Figure 174**
 182. **Figure 175**
 183. **Figure 176**
 184. **Figure 177**
 185. **Figure 178**
 186. **Figure 179**
 187. **Figure 180**
 188. **Figure 181**
 189. **Figure 182**
 190. **Figure 183**
 191. **Figure 184**
 192. **Figure 185**
 193. **Figure 186**
 194. **Figure 187**
 195. **Figure 188**
 196. **Figure 189**
 197. **Figure 190**
 198. **Figure 191**
 199. **Figure 192**
 200. **Figure 193**
 201. **Figure 194**
 202. **Figure 195**
 203. **Figure 196**
 204. **Figure 197**
 205. **Figure 198**
 206. **Figure 199**
 207. **Figure 200**
 208. **Figure 201**
 209. **Figure 202**
 210. **Figure 203**
 211. **Figure 204**
 212. **Figure 205**
 213. **Figure 206**
 214. **Figure 207**
 215. **Figure 208**
 216. **Figure 209**
 217. **Figure 210**

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Amidation site.

amino acids 145-149

FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGAGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTC
AAGCCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA
GGTTGAGGAAC TGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG
CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTAT
CACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTTCAGAGACAGTTATCACATCTG
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC
ATGTACTGCCAGTTTGCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA
TGGCCACCAGGGGCGAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG
TGCTGTGCCTTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA
GCTTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG
GAGCCTTGAGCCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT
GCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA
TCTTCTTCCAGTAAGTTTCCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC
TCCCCAGGCTGTTCTCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAAAC
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTTGGAGGGGAGGAGATGGAAACAATGTGG
AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAA
CCTGGCAAAAATGCAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTTACCCTGCATTACATGTGTTTATTTCATCC
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTTCATATCCAAGATCAATTC
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCAATTGTTCTCCTCGTCCATCAGGGATCTCAGAG
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCACCAA
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG
GTCAAAC TAATTCTCACATCCCTCTAAAAGTAACTACTGTTAGGAACAGCAGTGTTCTCAC
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT
TGCATTAGTAAC TTTGAAAGGTATATGACTGAGCGTAGCATAACAGGTTAACCTGCAGAAACA
GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC
TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTTCAGGTGTCA
TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA
TAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 84

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRS AVEEMEAE EEA AAKASSEVNLANLP PSYHNETNTDTKVGNN TIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 85

AAGGAGGCTGGGAGGAAAGAGGTAAGAAAGGTTAGAGAACCTACCTCACATCTCTCTGGGCTCAGAAGGACTCTG
AAGATAACAATAATTTTCAGCCCATCCACTCTCCTTCCCTCCCAAACACACATGTGCATGTACACACACACATACA
CACACATACACCTTCCTCTCCTTCACTGAAGACTCACAGTCACTCACTCTGTGAGCAGGTATAGAAAAGGACAC
TAAAGCCTTAAGGACAGGCCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAACATGGGAGGGG
CCAGGCACGGTGACTCACACCTGTAATCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTGAGGAG
TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC
AGGTGCCTGTAATCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT
CAGCTGAGTGACCCGCTGCACTCCAGCCTGGGTGACAGAATGAGACTCTGTCTCAAACAAACAAACACGGGAGGA
GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACTCTGCATCCTCTTCTTCCAGGGCTGCCCCGTATGGGGCCTG
GCAATGACTGAGCAGGCCAGCCCCAGAGGACAAGGAAGAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG
GTGTAGAATGACTGCCCTGGGAGGGTGTTTCTTGGGCCCTGGCAGGGTTGCTGACCTTTACCCTGCAAAACACA
AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACCATGAGGCTTCTCGTGGCCCC
ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCCTCA
GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCCGCTCGTCTACCGCGAGGCTACCCTGTGGACTGCAATGA
CCTATTCTTGACGGCAGTCCCCCGGCACTCCCCGAGGCACACAGACCCTGCTCCTGCAGAGCAACAGCATTGT
CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCAGAACAGCTTTTCGGA
TGCCCGAGACTGTGATTTCCATGCCCTGCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT
GGAGGACCACAGCTTTGACAGGGCTGGCCAGCCTACAGGAACCTATCTCAACCACAACCAGCTCTACCGCATCGC
CCCCAGGGCCTTTTCTGGCCTCAGCAACTTGCTGCGGCTGCACCTCAACTCCAACCTCCTGAGGGCCATTGACAG
CCGCTGGTTTGAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT
GAACCTCCGGCCCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC
CCTGGAGGGGCTGCAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCAGGCGGGCACT
GGAACAGGTGCCCGGGCTCAAGTTCCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGGACTTTGC
CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT
GAACCTCCCCGAGCTGACCAAGCTGGACATACCAATAACCCACGGCTGTCTTCATCCACCCCCGCGCCTTCCA
CCACCTGCCCCAGATGGAGACCTCATGCTCAACAACAACGCTCTCAGTGCCCTTGCAACCAGCAGACGGTGGAGTC
CCTGCCCAACCTGCAGGAGGTAGGTCTCCACGGCAACCCCATCCGCTGTGACTGTGTATCCGCTGGGCCAATGC
CACGGGCACCCGTGTCCTTCATCGAGCCGCAATCCACCTTGTGTGCGGAGCCTCCGGACCTCCAGCGCCTCCC
GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCCTCATCTCCCCACGAAGCTTCCCCCAAG
CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG
GGTCACTCCAGCTGGGCTTCGACTGACACCTGCCCATGCAGGCAGGAGGTACCGGGTGTACCCCGAGGGGACCCT
GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCCAAGACCTGGTGGGGGCTGACAC
TAAGACGGTTAGTGTGGTTGTGGGCCGTGCTCTCCTCCAGCCAGGCAGGGACGAAGGACAGGGGCTGGAGCTCCG
GGTGAGGAGACCCACCCCTATCACATCCTGTCTATCTTGGGTCACCCCACCAACACAGTGTCCACCAACCTCAC
CTGGTCCAGTGCCCTCCTCCCTCCGGGGCCAGGGGGCCACAGCTCTGGCCCGCTGCTCGGGGAACCCACAGCTA
CAACATTACCCGCTCCTTCAGGCCACGGAGTACTGGGCTGCCTGCAAGTGGCCTTTGCTGATGCCACACCCA
GTTGGCTTGTTATGGGCCAGGACCAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCTTGGGCTCAT
TGCCATCCTGGCTCTCGCTGTCTTCTCCTGGCAGCTGGGCTAGCGGCCACCTTGGCACAGGCCAACCCAGGAA
GGGTGTGGGTGGGAGGCGGCCCTCCTCCAGCCTGGGCTTCTTGGGGCTGGAGTGCCCCCTTCTGTCCGGGTTGT
GTCTGTCCCCCTCGTCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC
ACCATTGTCTCAAAATCTTGAAGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTTACCAAA
AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGC
CAAGACAGATGGGGCTTTGTGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA
CCTCTGCTGCCATTCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCTCCCCATCTT
CTCTCTGCCAGAGGCTCCTGGGCCCTGGCTTGCTGTCCCTACCTGTGTCCCGGGCTGCACCCCTTCTCTTC
TCTTTCTCTGTACAGTCTCAGTTGCTTGTCTTGTGCTCCTGGGCAAGGGCTGAAGGAGGCCACTCCATCTCAC
CTCGGGGGGCTGCCCTCAATGTGGGAGTGACCCAGCCAGATCTGAAGGACATTTGGGAGAGGGATGCCAGGAA
CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTCTATAGGCAATTTTGTACCTTGTGGAGAA
ATGTGTACCTCCCCAACCCGATTCACTCTTTTCTCCTGTTTGTAAAAAATAAAAAATAAATAAATAA
AAAA

[illegible]

amino acids 470-474, 660-664, 692-696

FIGURE 87

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGACATTGTGTACCGCCT
CTACATGCGGCGAGACCATCATCAAGGTGATCAAGTTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA
CATCAAGTTTCGACGTGGACTGCACCGTGGACATTTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCCACCC
CCTGGCCACACTCTTCAAGATCCTGGCGTCTCTTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG
CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA
GCGCTTCGCGCTCTTCTGTGCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT
GGACAAGCTCCGGCAGCGGCTCACCAGAAGCGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCCGCC
CAGCATTGCCCGAGCTCACGGGCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT
GGCCTTCTGCGCGAGAACCCTGCGGGCGCTGCACATCAAGTTCAACCGACATCAAGGAGATCCCGCTGTGGATCTA
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA
CGGGCTGCGGGAGCTCAAACGCCCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCAC
AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAA
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCAGACCT
GCACCGCCTCACCTGCCCTTAAGCTGTGGTACAAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACAA
CCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGAAGCTGCG
CTACCTGGACCTCAGCCACAACAACCTGACCTTCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT
AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG
CAACCGGCTGGAGTGCTGCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCCTGAGCGAG
GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCGGAGGGGCGAGGCTAGCTTCTCCAG
AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGA
GGACAGTATCTGTGGGGCTGGCCCCCTTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG
CTGCCACCAGAGGTCTGGGACCTCACTTTAGTTCTTGGTATTTATTTTCTCCATCTCCACCTCCTTCATCC
AGATAACTTATACATTCCCAAGAAAGTTAGCCCGAGTGAAGGTGTTTCAAGGAAAGGTGGGCTGCCCTTTTCCCC
TTGTCTTTATTTAGCGATGCCGCCGGGCATTTAAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCAG
CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCTCCAGCTGGA
AAGGCCAGGCCTGGAGCTTGCTCTTCAGTTTTTGTGGCAGTTTTAGTTTTTTGTTTTTTTTTTTAAATCAAA
AAACAATTTTTTTTTAAAAAAGCTTTGAAAATGGATGGTTTGGGTATTAAGGAGAAAAAAGAAAAAAGCTTAAAAA
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGACGT
TGAAGTGTGTTTTCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGGATCTGGTGTGACCTTGGTCCAGGAGTT
CTATTTGTTCTCTGGGAGGGAGGTTTTTTTGTGTTTTTTTGGGTTTTTTTGGTGTCTTGTGTTTTCTTCTCCTCC
ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG
ACTCGGGTTGGCTAATCCCCGGATGAACGGTGTCTCATTCGCACCTCCCCCTCCTCGTGCCTGCCCTGCCTCTCCA
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTCCCCACCTCCTGCGGCATGGGTGTGT
CCAGTGCCACCGCTGGCCTCCGTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTTCATGAAGAGCAGACACTTA
GAGGCTGGTCCGGAATGGGGAGGTGCGCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGC
CTGGAGTGACACAGCCAGTGGGACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT
AGAAGGGTCCCCGCCTTAGATCAATCAGTGGACACTAAGGCACGTTTAGAGTCTCTTGTCTTAATGATTATGT
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTGTGTCATTGGATATAATCCTCAGAAAATAATGCACACTAG
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTGTAGACTCGGTACAGTATCAAATAAA
ATCTATAACAGAAAAA

FIGURE 88

MRQTIKVIKFILIIICYTVYYVHNIFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQLTKNAQDKLELHLFMLS GIPD TVFDLVELEV
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRALHIKFTDIKEIPLWI
YSLKTLLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI
NNEGTKLIVLNSLK KMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIIISFQ
HLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIP TQLFYCRKLRYLDLSHNNLT
FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRLHLGNNVLQSLPSRVGELTNLTQIE
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

Transmembrane domain:

amino acids 51-75 (type II)

N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,
398-402, 493-497

N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCAATGGAGCTGGCACTGCGGCGCTCTCCCGT
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT
GGCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCCTGGTCAT
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTTGAAACTTTGAGGAAATTGGGC
CCCTTGACAGTGATCTCAAACCACGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT
GTGGATAATCCCGTGGGCACTGGGTTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG
AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC
ATTGGTCTAGAGCTTTATAAGGCCATTACGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT
TGCCTTGGGTGATTCTCGGATCTCCCTGTTGATTCTGGTGCTCTCCTGGGGACCTTACCTGT
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT
GATCATTGAACAGAACACAGATGGGGTGAACCTTCTATAACATCTTAACTAAAAGCACTCCCA
CGTCTACAATGGAGTCGAGCTAGAATTACACAGAGCCACCTAGTTTGTCTTTGTCTCAGCGC
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC
AACGTGACGGTGATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTGAGGAGGCCTG
GGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTCAGTCAGCTGAAGTGAAGGCCCTGT
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC
TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT
GATGAGACTGGTGACTCAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT
TGGGGCACAGAGCTGAGCTGAGGCCGTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT
AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAACCTAAGATTTTAAAAAATTGAT
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

FIGURE 90

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSLKPRKTTWLQAASLLFVDNPVGTGFSY
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNGLYRE
ATELWGKAEMII EQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQORDALS
QLMNGPIRKCLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL
IVDTMGQEAWVRKLKWPPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP
SDQGDMAKMMRLVTQQE

Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,
187-193, 195-201, 331-337, 332-338, 360-366

FIGURE 91

GGCCGCGGGAGAGGAGGCCATGGGCGCGCGCGGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCA
GGGGAGCCTGCGCCTGTGGGATTCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG
T TACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG
CCTTGGTGAAGCTGTCTGCACCTGTACCTACACTAAACACATCCAGCCCATCTGTCTCCAG
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA
AGAGGATGAGGCACCTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA
ACTCTATGTGCAACCACCTCTTCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG
GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC
CACTGCCAAGTCAGGCCCTGGTTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

"GTCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA"

FIGURE 92

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSRLRW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDI FGDMVCAGNAQGGKDACFGDSSGGPLACNKNG
LWYQIGVVSWSVGCGRPNRPGVYTNISHHFEWIKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 167-171, 200-204, 273-277

Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,
259-265, 269-275

Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAAATGGGACTCCAAGCCTGCCTCCTAGGGCT
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC
TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCCGATCCCAG
CTCTCCTCAATACGGAAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC
CACTGACCCTCCACACGGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT
GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC
TGGGGCTGAGTTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC
ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT
TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT
GCATCTGGGGGTAACCCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTTCTGGAGCAGTATTTCCATGAC
TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCGGGATTGAGGCCAGTCTAGATGTGCAGT
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG
GGACAGGAGCCCTTCTGTCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCCTCAGGTGACAGT
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG
CCCCATGTCAACACAGTGGGAGGCACATCCTTCCAGGAACCTTTCTCATCACAAATGAAA
TTGTTGACTATATCAGTGGTGGTGGCTTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG
GAAGCTGTAACGAAGTTCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACCTTTCTGATGGCTACTGGGTGGTCAGCAACA
GAGTGCCCATTCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTGGGGGGATCCTA
TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCTCTTGGCTTTCTCAACCCAAG
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCTGGCTGGGATCCTGTAACAGGC
TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC
AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG
GAAGCCCTGCTGAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT
CCTGCCCTTAGATTCTCAATAAGATGCTGTAAGTAGCATTTTTTTGAATGCCTCTCCCTCCGC
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA
TTTCACTTGATATTCAATCCCCAATTCAGTCAAGGAGACCTCTACTGTCACCGTTTACTCT
TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACCTTCTCTG
ACTACTCTTGTCTTCTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTTGTCCCCTGGAACAAATCACTGACA
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA
TGTAACAAAA

FIGURE 94

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

Signal sequence:

amino acids 1-16

N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,
521-527, 533-539, 549-555

[illegible][illegible]

FIGURE 96

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKS
RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPGWIKGNANDIGMDYDYA
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD
AQPASGSGVYVRMWKRQQQKWERKIIGIFSGHQWDMNGSPQDFNVAVRITPLKYAQICYW
IKGNYLDCREG

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

bioRxiv preprint doi: <https://doi.org/10.1101/150000>; this version posted May 1, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 97

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGGTGGTTT
CTGGAGCGCCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC
ACTGCTGCCCCACTGTTTCAAGGACAACCTGAACAAACCATACTGTTCTCTGTGCTGCTGGG
GGCCTGGCAGCTGGGGAACCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC
CCCACCCTGTGTATTCTTGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG
CGCTCCATACAGTTCTCAGAGCGGGTCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCTTG
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCGCCGTAAATAGGCTCATCTACC
TCTACCTCTGGGGGCCCCGGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCCGAC
GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCAACGGCCTCATGTCCCCGCCCCAC
GACTTCCGGCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT
AGGTATTTGTAACCCTGCCCACATATCTTATTTATTCCTCCAATTTCAATAAATTATTTATT
CTCCAAAAAAAAAA

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVMSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP
WIVSIQKNGTHHCAGSLLTSRWVITAACHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPNTHCWISGWGSIQDG
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC
QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG
SGAAARS

Signal sequence:

amino acids 1-32

N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

FIGURE 99

GACGGCTGGCCACCATGCACGGCTCCTGCAGTTTCTGATGCTTCTGCTGCCGCTACTGCTA
CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC
AACAAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTGAGGAGACCAACATCGAATTACT
GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAAACCAGGCC
CCAACCTTCCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA
CCTTCCCCAAATCGACCCATGTTTCTATCCCCAAATCAGCAGACAAAGTGACAGACAAAACA
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG
GGAACCTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACA
AGCCTAGCGTTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG
TGAAGAGGTCAGCTGTCCTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCATAACAAGATA
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACC
GGGCCCACACCTCTCCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT
CACTGCCTACCTGGCCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG
TGTAGCTGGGGATGGGGATTCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT
GTCCAACCTGTGAAA

FIGURE 100

MHGSCSFLMLLLPLLLLLLVATTGPVGALTDEEKRLMVELHNLRYAQVSP TASDMLHMRWDEE
LAAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ
MCGHYTQVVWAKTERIGCGSHFCEKLOGVEETNIELLCNYEPPGNVKGKRPYQEGTPCSQC
PSGYHCKNSLCEPIGSPEDAQDLPLYLVTEAPSFRATEASDSRKMGT PSSLATGIPAFVLVTEV
SGSLATKALPAVETQAPTSLATKDPSPMATEAPPCVTTEVP SILAAHSLPSLDEEPVTFPKS
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAEELPPSSEVLAS
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV
SGLNSGPGHVWGPLLGLLLLPLPLVLGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,
250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 101

GTAAGTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTTCGGTCATTCTCCAAGTTATGGTGGACGT
 ACTTCTGTTGTTCTCCCTCTGCTTGTCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTTCAT
 CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAAGTGAAACTGAACAACAATGAATTGGAGACCATTCC
 AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA
 ACATCTGAAAGAGTTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAAACTGCATT
 TCCAGCCCCACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA
 TTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCAAGATGTTTAAACT
 GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG
 TGCTCTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTTGGGGGCTGAGCAA
 CATGGAAATTTTGCAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCT
 GCAGGAACCTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG
 TGAGCTGGACCTAACTTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATAC
 ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT
 GGATCTGAAGAACAATGAAATTTCTGGACTATTGAAGACATGAATGGTGTCTTTCTCTGGGCTTGACAAACTGAG
 GCGACTGATACTCCAAGGAAATCGGATCCGTCTTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA
 TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCAAAATGAAGAACTGCAACAATT
 GCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAAATGGCTCCACAGTGGGTGGCGGAAAACAACCTT
 TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAAGGAAGAAGCATTTTGTCTGTTAGCCCAGA
 TGGCTTTGTGTGTGATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAAACAGTCGGCAATAAAAGGTTTC
 CAATTTGAGTTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGTCTGGAAAAAAGACAATGA
 ACTACTGCATGATGCTGAAATGGAAAAATATGCACACCTCCGGGCCCCAAGGTGGCGAGGTGATGGAGTATACCAC
 CATCCTTCCGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCTCCAATCACTTTGGTTTC
 ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCAT
 CCGAGCTGGGGCCATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCAGCCCCCAGATAGCCTGGCAGAAGGA
 TGGGGGCACAGACTTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGT
 GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAGCAAATGC
 AACTCTGACTGTCTTAGAAACACCATCATTTTTCGGGCCACTGTTGGACCGAAGTGAACCAAGGGAGAAACAGC
 CGTCTTACAGTGCAATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGT
 AACCGAGAGGCACTTTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTGATGCTGGGAA
 ATACACATGTGAGATGCTTAACACCCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC
 CTGCGACTCCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTGCTGATCATAGC
 CGTGGTTTGTGTGTGGTGGGCACGTCACTCGTGTGGGTGGTTCATCATATACCACACAAGGCGGAGGAATGAAGA
 TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCTTAGTTATTTGTCTCATCTCAGGGAACGTT
 AGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTTCACATCTTCAGGTGCTGG
 ATTTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC
 CACAGATCTGTTCTTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA
 TCCTTTTGAACATATCATACAGGTTGCAGTCTTGACCAAGAACAGTTTAAATGGACCACTATGAGCCCAGTTA
 CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC
 GTGGCCTTCACATGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTG
 TCTAAACAAGTCTCTTTAGATTTTAGTGCAATCCAGAGCCAGCGTCGGTTGCCCTCGAGTAATCTTTTCATGGG
 TACCTTTTGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTTCAGCC
 AAGAGCCTTTTATTTGAAAGCTCATTTCTCCCCAGACTTGGACTCTGGGTTCAGAGGAAGATGGGAAAGAAAGGAC
 AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTC
 TTATGACTTGGACACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACCTTTTATTTA
 AAAGAGAGAGAATCTTATGTTTTTTTAAATGGAGTTATGAATTTTAAAAAGGATAAAAAATGCTTTATTTATACAGAT
 GAACCAAAATTAACAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTAAACTA
 TTTTTTAACCTTTGTTTTATGCAAAAAAGTATCTTACGTAATTAATGATATAAATCATGATTATTTTATGTATTT
 TTATAATGCCAGATTTCTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT
 TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

bioRxiv preprint doi: <https://doi.org/10.1101/030660>; this version posted April 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 102

MVDVLLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVVSAN
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP
GYFDNLANTLLVLKLNRRNRI SAIPPKMFKL PQLQHLELNRNKIKNVDGLTFQGLGALKSLKM
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDWE
FCQKLSELDLTFNHL SRLDDSSFLGLSLNLTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE
ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ
MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENYHLRAQG
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA
MARLECAAVGHPAPQIAWQKDGGTD FPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN
SAGSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPKLNWTKDSDPLVVT
HFFAAGNQLLIIVDSVDSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG
WATVGVV IIAVCCVVGTSLVWVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGT
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP
MYLKGNVYGSDPFETYHTGCSDPDRTVLMDHYEPSYIKKKECYPCSHPSSESCERSFSN
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDA
YSSFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEEHICTFKQTLNRYRTPNFQS
YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,
1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,
612-618, 623-629, 714-720, 873-879

Summary

FIGURE 104

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLNLHENQIQIIVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYNLAMLNREIPNLTPLIKDELDSLGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNDIL
WLSWWIKDAPSNNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASSTLSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTQDTGMYTCMVSNSVGN
TTASATLNVTAATTTTFFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKFTTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVTNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,
442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGCA
TGAAAGGGCTTCGCCGCCGGGAGTAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC
GAGGGCGGGCGTGACCCCTCGGCTGGAAGTTTGTGCCGGGCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGA
GACCTAGGCGCGTGGACCGCGATGAGCGCGCGAGCCCTCCGTGCGCGCGCCGCGGGTGGGGCTGCTGCTGTGC
GCGGTGCTGGGGCGCGTGGCCGGTCCGACAGCGCGCGTCCGCGGGGAACCTCGGGCAGCCCTCTGGGGTAGCCGCC
GAGCGCCCATGCCCCACTACCTGCCGCTGCCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT
CCCGAGCCACTCCCGTCTCGGTGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTTCATCAAGGCAAGTTCC
ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACCTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA
GTCTCGGCAAATATTACACTTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG
TTTCAGTCCCTTGAACTTTGGACCTTAGCAGCAACAATATTTAGAGCTCCAACTGCATTTCCAGCCCTACAG
CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCACACACA
CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCACCCCAAGATGTTTAAACTGCCCCAACTGCAA
CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTTCAAAGGCCTTGGTGCTCTGAAGTCT
CTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG
CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTTCAT
CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA
ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCCTGGCCTAAGCTTACTAAATACACTGCATTTGGG
AACAACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC
AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC
CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTTGGAGCATCTAGACCTGAGT
GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCAAAAATGAAGAAACTGCAACAATTCATTTAAATACA
TCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAAACTTTTCAGAGCTTTGTA
AATGCCAGTTGTGCCCATCTCAGCTGCTAAAAGGAAGAAGCAATTTTGTGCTGTAGCCAGATGGCTTTGTGTGT
GATGATTTTCCCAAACCCAGATCAGGTTTCCAGCAGAAACACAGTCCGCAATAAAAGGTTCCAATTTGAGTTTC
ATCTGCTCAGCTGCCAGCAGCAGTATCCCCAATGACTTTTGTCTGGAAAAAGACAATGAACTACTGCATGAT
GCTGAAATGGAAATTTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAACCATCTTCGGCTG
CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCAATCACTTTGGTTTCATCCTACTCTGTC
AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC
ATGGCAGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCAGATAGCTTGGCAGAAGGATGGGGGCACAGAC
TTCCAGCTGCACGGGAGAGACGCATGCTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA
GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTAGCAAATGCAACTCTGACTGTC
CTAGAAACACCATCATTTTTCGGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCTACAGTGC
ATTGCTGGAGGAAGCCCTCCCCCTAACTGAACTGGACCAAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC
TTTTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATCTTTTCATGGGTACCTTTGGAAAA
GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTGAGCCAAAGAGCCTTTAT
TTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA
GAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC
ACATAGACTGAATGAGACCAAAGGAAAAAGCTTAACATACTACCTCAAGTGAACCTTTATTTAAAGAGAGAGAAT
CTTATGTTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAAATGCTTTATTTATACAGATGAACCAAAATTAC
AAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTTAACTATTTTTTAACTTTG
TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTTTATAATGCCAGA
TTTTCTTTTTATGGAATAAGTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT
ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAA

FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELGQPSGVAAERPCPTTCRCLGDLDDCSR
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNELETIPNLGPVSANIT
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY
FDNLANTLLVLKLNRRNRI SAIPPKMFKLPLQLHLELNRNKIKNVDGLTFQGLGALKSLKMQR
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDWEFC
QKLSELDLTFNHLRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS
WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK
KLQQLHLNTSSLLCDCQLKWLPOQVAENNFSQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDE
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENYAHLEAQQGE
VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA
RLECAAVGHPAPQIAWQKDGDTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA
GSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF
FAAGNQLLIIVDSVDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA
TVGVVVIIVVCCVVGTSVLVWVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTADLRQ
DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY
LKGNVYGSDDPFETYHTGCSPPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS
HVRKLLNTSYSHNEGPGMKNLCLNKSSLDIFSANPEPASVASSNSFMGTFGKALRRPHLDAYS
SFGQPSDCQPRAFYLKAHSSPDLDGSEEDGKERTDFQEEHICTFKQTLNRYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,
1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

FIGURE 107

CAAAACCTTGCCTCGCGGAGAGCGCCAGCTTGAATTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC
TGGGGGAGCGCGTTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC
CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCTGGGTGGTCCCGTCCCTATCCCTCCTTTATATA
GAAACCTTCCACACTGGGAAGGCAGCGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG
GCGCACAGCATTCGAGTTTACAGATTTTACAGATACCAAATGGAAGGCGAGGAGGCAGAACAGCCTGCCTGGT
TCCATCAGCCCTGGCGCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCAGAGCCGGGTGCTGC
TGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA
GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGTGTTACTGAGCCCTGAGG
AGCCCGGGCCTGGCCAGCCGCGGTGAGTGCCTGAGACTGTGCTGTTCCAGGAGGGCGTCTGTTGAGTGTG
GCGGTATTGACCTGCGTGAGTTCCCGGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACCAACAGC
TGGAAGAGATCTACCTGAGGAGCTCTCCCGGTGCACCGGCTGGAGACACTGAACCTGCAAAACAACCGCCTGA
CTTCCCGAGGGCTCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAAATAACAAGC
TGACCTTGGCACCCCGCTTCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT
ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC
TGCCGGAACAACATGTTCACCGCTCCAGCAACGTGAGGTCTCATCTGTCCAGCAACTTCTGCGCCACGTGC
CCAAGCACCTGCCGCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT
TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA
CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCAGCTGGGCTGC
CGCGCAGCCTGGTGTGCTGCTGCACTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCATCC
GCAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCCACTGGCCTTCCAGGGCC
TCAAGCGGTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCGTGCCAGTGGCCTGCCTCGCCGCGTGC
GCACCTCATGATCTTGCAACAACAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTTGGAGG
AGCTCAACCTCAGCTACAACCGCATCACAGCCACAGGTGCACCGCGACGCTTCCGCAAGCTGCGCCTGCTGC
GCTCCTGGACCTGTGCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG
TCAAGCGCAATGAGCTGGCTGCCTTGGCAGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA
CCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA
TCGCCGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACACA
AGATTAGTGCGGTGCCCGCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAAACAAGC
TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCACT
TAGAGTTTGGTGACATTTCCAAGGACCGTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAGGAAGAGGAGGAGGAGG
AGGAAGAGGAAAACAAGATAGTGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGGACTCTTTTCTGC
AGCACACGCTGTGTGCTGTGAGCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA
TCCACATGACACGGCTGACACAGTCTCATATCCCCACCCCTTCCACGGCGTGTCCACGGCCAGACACATGC
ACACACATCACACCTCAAACACCCAGCTCAGCCACACACAACCTACCTCCAAACCACCACAGTCTCTGTACAC
CCCCACTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGTCTGCCCTGCCCTGGCACACACAGGCACCCA
TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACCACACACACACATGCACAAGTCATGTGCGAA
CAGCCCTCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT
GTCCATCTGTCCGTCCGTTCCTTGGAGAAGACACAAGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCCTCT
GGAACCTCAAAAAGCTGGCTTTTATTTCTTTCCCATCTATGGGGACAGGAGCCTTCAAGGACTGCTGGCCTGGCC
TGGCCACCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA
CAGGCACTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG
CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAGGAAGGACCCAGCTGCACCTAGGAGACACCTT
GTTCTTCAAGCCTGTGGGGGAAGTTCCGGGTGCCTTTATTTTTTATTTCTTTTCTAAGGAAAAAATGATAAAAA
CTCAAAGCTGATTTTTCTTGTATAGAAAACTAATATAAAGCATTATCCCTATCCCTGCAAAAAA

FIGURE 108

MEGEEAEQPAWFHQPWPRPGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSP EENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFFPGDLP
EHTNHL SLQNNQLEKIYPEELSRLHRLETNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSV DANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTL MILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLRLSLDL SGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSSVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGR LGKEKEEEEEEEEEEEETR

Signal sequence:

amino acids 1-48

N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,
477-483, 498-502, 539-545, 548-554

Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,
535-557

[illegible]

GGAGGGGGGCTCCGGGCGCCGCGCAGACAGACCTGCTCCGGGCGCGCGCCTCGCCGCTGTCTTCCGGGAGCGGCAG
CAGTAGCCCGGGCGGCGAGGGCTGGGGTTCTCTGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCACCACCC
CAACCTGTTTCTCGCGGCCACTGCGCTGCGCCCCAGGACCCGCTGCCAACATGGATTCTCTCTGGCGCTGGT
GCTGGTATCTCTCGCTCTACCTGCAGGCGGGCCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTTCATCGAT
TGGCTATGTCTGTTATGGTGGGAGGATTGACTGTCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTACGCC
TGTGTGCCAACACAGATGCAAAACATGGTGAATGTATCGGGCCAAACAAGTGCAGGTGCATCTGTTATGTCTGG
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAAGCAGGTGCATGAACACTT
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCCTGCTCAAGTGCCCTGACCTG
CTCCATGGCAAACCTGTGATGCTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCCTGGCCT
GCACCTGGCTCTGTATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCCTCCTGCCCTAGATT
TAGGCAATGTGTCAACACTTTTGGGAGCTACACTGTCAAGTGTCTATAAAGGCTTCGATCTCATGTATATTGGAGG
CAAATATCAATGTCTAGCATAGACGAATGCTCACTGGTTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTTATAA
CGTACGTGGGTCTTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGTGGAGTGCATGCTGTTATGTTTATAA
AGTTATGATTGAACCTTCAGGTCCAATTTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAA
TAATAATTGGATTCTGTATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATCTCTCTATCATTACCAA
CAGGCTACTTCTAAGCCCAACAAGACCTACACCAAAGCCCAACCAATCTCTACTCCACCACCACCACCACC
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACCCGACTGACAACTATAGC
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCTCAGAAACCCAGG
AGATGTGTTTCAGTGTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTGGCAGCCAAAGCCCCAGG
GGGAAAGCTGCACGCTTGGTGCTACCTCTCGGCCCGCTCATGCATTACGGGGACCTGTGCTGTCAATTCAGGCA
CAAGGTGACGGGGCTGCATCTTGGCAGCTCCAGGTGTTGTGAGAAAAACCGGTGCCACGGAGCAGCCCTGTG
GGGAAGAAATGGTGGCCATGGCTGGAGGCAACACAGATCAGTCTTGCAGGGCTGACATCAAGAGCGAATCACA
AAGATGATTAAAGGGTTGGAaaaaaagATCTATGATGGAAAAATTAAGGAACTGGGATTATTGACCTGGAGAAG
AGAAGACTGAGGGGCAAACCAATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG
TTCTCCATATGCACATAAGAAATAGAAACAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG
GGCCATTGTTAGAACTACTCATAaaaaaagAGGTGTGAAAACTCAGTATCTCTCTCTTTCTAAAAAATTAGA
TAAAAATTGTCTATTTAAGATGGTTAAAGATGTTCTTACCAAGGAAAGTAAACAAATTAGAAATTTCCCAA
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAACTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG
CATTTCCCTCTTGACCTCTTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAATGCTGAGCTCACTGAAATA
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATTCCAAATGAGAGTATGATGGAC
AGATATTTTAGTATCTCAGTAATGTCTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTCATGGTAAAGGTATAAGCC
TTTCATTTGTTCAATGGATAGTGTTCAGATTTTTTTTTTTTAAAGAGATCCTTCAAGGAACAGTTTCAGAGAG
ATTTTCATCGGGTGCAATCTCTCTGCTTCGTGTGTGACAAGTTATCTGGCTGCTGAGAAAGAGTGCCTGCCCC
ACACCGGCAGACCTTTCTTCACCTCATCAGTATGATTAGTTTTCTCTTATCAATTGGACTCTCCAGGTTCCAC
AGAAGCAGTAATATTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTAACCTGGTAAAGGCAGGGCTGG
AGGGGGAAAAATAAATCAAGCCTTTGAGTAAACGGCAGAATATAGGCTGTAGATCCATTTTTAATGGTTTCATT
TCCTTTATGGTCAATAAATGCACAGCTGAAGATGAAGGGGAAAAATAAATGAAATTTTACTTTTTTCGATGCCAA
TGATACATTGCACTAAACTGATGGAAGAAAGTTATCCAAAGTACTGTATAACATCTGTTTTATTATTATGTTTTT
CTAAAAATAAAAAATGTTAGTGGTTTTCCAATGGCCTAATAAAAACAATTATTTGTAAATAAAAAACTGTTAGTAAT

FIGURE 110

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP
RCKHGEICIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD
GSCSSALTCSMANCQYGCDVVKGQIRCQCPSGLHLAPDGRTCDVDDECATGRASCPRFRQC
VNTFGSYICKCHKGFDLMIYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQG
DGLTCVYIPKVMIEPSGPIHVPKGNGTILKGD TGNNNWIPDVGSTWWPPKTPYIPPIITNRP
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDN
RVQTDPOKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG
KAARLVLPGLRLMHSGDLCLSFHVKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI
TLRGADIKSESQR

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,
421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCTTTAGATTGTGA
AATGTGGCTCAAGGTCTTCACAACTTTCTTTCTTTTGAACAGGTGCTTGCTCGGGGCTGA
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTGAGGCCCTCTACCTACCCGTC
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC
AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCT
GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACCTATCTGCCAGTCA
GAAGATACAAGTCACGGTTGATGATCCTGTCAAAAGCCAGTGGTGCAGATTCATCCTCCCT
CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC
TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT
GCCTGGTGAGGAACCCTGTGAGTGAATGGAAGTGATATCATTATGCCCATCATATATTAT
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT
GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC
CGGCAGGCAAGATGAAACTCATTTACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG
CACAGAAAGGAAAATCATTGTACCTTTAGCAAGTATAAAGTGAATATCACTATTTTTGATT
ATATCCATGTGTCTTCTCTTCTATGGAAAAATATCAACCCTACAAAGTTATAAAACAGAA
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT
GTATGAAGTTATTTCAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACCTTCATGG
GCTAAACAGTACATTTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG
GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA
ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA
TGTGGAAACTTTACATTGTTTCGATTTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG
TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA
CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACCTATGCCTTCTCTTTTTTTTCAATCACC
AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA
TTTTTTTTCAAGGAAAGATGGATTCAAATAAATTATCTGTTTTTGCTTTTAAAAAAAAAAAAA

FIGURE 112

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQNGTILSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,
276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,
239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

[illegible]

GCAAGCGCGCGAAATGGCGCCCTCCGGGAGTCTTGCAGTTCCCTTGGCAGTCCTGGTGCTGTT
GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA
ACTGGAGAGAACTGTCTGGGAAGGAGACTGGATGATAGAATTTTATGCCCGTGGTGCCCTGCT
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA
TATTGCGAAAGTAGATGTACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGAAGAGTATTGAGCCCGTTTCATCATG
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT
GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC
AGATTGCCTTTGTCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCCTTCAAAAAAAT
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGC
CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG
TTATCTTAATATTATGATTTTGATAAAAACAGAAGATTGATCATTTTGTTTGGTTTGAAGTG
AACTGTGACTTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCAATTAAATTGAAGAGTCTA
CATTACAGAACATAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG
GTTTAAATAGTTCCTCTAATTTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT
TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG
TTTCACTGTGTGAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG
AAATGTGTATTTTCAGTGACAATTTTCGTGGTCTTTTTTAGAGGTATATTCAAAATTTCTTGT
ATTTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA
TGGTAATACAGGATATGCTACTGATTTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC
CAACAAAGTTTGATTTTTCTCTTGATTTTTCTTACTTACTATGGGTTACATTTTTTATTTTT
CAAATTGGATGATAATTTCTTGGAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTTGT
GTTTCAAACCTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAATTT
TTGGCCACTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT
TTTCTTTTTTGGATGTGAAGGTGAACATTTCTGATTTTTTGTCTGATGTGAAAAGCCTTGGA
TTTTACATTTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG
CATCTTCTTGATATATGTCTTAAATGTATTTTTTGTCTCATATACAGAAAGTTCTTAATTGAT
TTTACAGTCTGTAATGCTTGATGTTTTAAATAATAACATTTTTTATATTTTTTAAAGACAA
ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTTCACAGGTAAAA
GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATAACATAGTTTT
CACCTTAAAGAAGGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC
AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT
AAAGAGTTTGGATGTGTAACCTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT
TTCTAACCACTTCATTAAGCTGAAAAAAAAAAAAAAAAAAAA

FIGURE 114

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL
QPEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI
NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL
ATLFSGLLLGLCMIFVADCLCPSKRRRRPQYPYPYPSKKLLSESAQPLKKVEEEQEAEDEEDVSE
EEAESKEGTNKDFPQNAIRQRS LGPSLATDKS

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

FIGURE 115

CGGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTAACTAATTCAACAAACGGGACCCTT
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG
AAGGTCTTGGCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAACTCTGGGTGTCTGC
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGG
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGTCAGGACTCAGCTGCA
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCAGGCACAATTTCCA
GAGGGGAACCTTCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATACGGTGTGTGGATGGAGCAGGTGCT
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA
GGAAGAGAGAGAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAAGGCA
TCTATGAGAGGCCTGTGGACTACCTGCCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGTTGG
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA
TCGTAGGTACTACGATGTCTGCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA
CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG
TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAGACATGCTGCCTGCCC
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT
TGAGACCTTGTGGATCAACAGAAGTTGACTTGACATCCTTTTCTGTCTTCCCCCTTCCCTGGTC
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTTCTTTGTATGTTTCTATCAGGCT
GATTTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT
GTGACTGAAGTCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTGTTTTAT
GGCATTTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

FIGURE 116

MKLWVSALLMAWFGVLSQVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA
NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFP
TDEDEIGAALKMRLQDTYRLDPGTISRGLPGTKYQAMLSVDDCFGMGRSAYNEGDYHTV
LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLR
YFEQLLEEEREKTLTNQTEAELATPEGIYERPDYLPERDVYESLCRGEQVGLTPRRQKRLF
CRYHHGNRAPQLLIAPFKEEDEWDSPIVRYVDVMSDEEIERIKEIAKPKLARATVRDPKTG
VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHDFS
RRPFDGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEQDYL
TRHAACPVLVGCKWVSNKWFHERGQEFRLPCGSTVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,
346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

Downloaded from www.sciencedirect.com

Country	Year	Population (millions)	Urban population (millions)	Urban population (%)	Population density (per sq km)	Urban population density (per sq km)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)
Algeria	1980	10.0	4.0	40.0	100	250	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1985	10.5	4.5	42.9	105	263	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1990	11.0	5.0	45.5	110	276	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1995	11.5	5.5	47.8	115	289	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2000	12.0	6.0	50.0	120	302	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2005	12.5	6.5	52.0	125	315	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2010	13.0	7.0	53.8	130	328	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2015	13.5	7.5	55.6	135	341	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2020	14.0	8.0	57.1	140	354	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2025	14.5	8.5	58.6	145	367	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2030	15.0	9.0	60.0	150	380	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2035	15.5	9.5	61.3	155	393	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2040	16.0	10.0	62.5	160	406	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2045	16.5	10.5	63.6	165	419	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2050	17.0	11.0	64.7	170	432	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2055	17.5	11.5	65.7	175	445	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2060	18.0	12.0	66.7	180	458	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2065	18.5	12.5	67.6	185	471	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2070	19.0	13.0	68.4	190	484	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2075	19.5	13.5	69.2	195	497	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2080	20.0	14.0	70.0	200	510	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2085	20.5	14.5	70.7	205	523	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2090	21.0	15.0	71.4	210	536	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2095	21.5	15.5	72.1	215	549	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2100	22.0	16.0	72.7	220	562	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2105	22.5	16.5	73.3	225	575	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2110	23.0	17.0	73.9	230	588	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2115	23.5	17.5	74.5	235	601	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2120	24.0	18.0	75.0	240	614	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2125	24.5	18.5	75.5	245	627	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2130	25.0	19.0	76.0	250	640	1.5	2.5	1.5	2.5	1.5	

Country	Year	Population (millions)	Urban population (millions)	Urban population (%)	Population density (per sq km)	Urban population density (per sq km)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)	Population growth rate (%)	Urban population growth rate (%)
Algeria	1980	10.0	4.0	40.0	100	250	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1985	10.5	4.5	42.9	105	263	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1990	11.0	5.0	45.5	110	276	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	1995	11.5	5.5	47.8	115	289	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2000	12.0	6.0	50.0	120	302	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2005	12.5	6.5	52.0	125	315	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2010	13.0	7.0	53.8	130	328	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2015	13.5	7.5	55.6	135	341	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2020	14.0	8.0	57.1	140	354	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2025	14.5	8.5	58.6	145	367	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2030	15.0	9.0	60.0	150	380	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2035	15.5	9.5	61.3	155	393	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2040	16.0	10.0	62.5	160	406	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2045	16.5	10.5	63.6	165	419	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2050	17.0	11.0	64.7	170	432	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2055	17.5	11.5	65.7	175	445	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2060	18.0	12.0	66.7	180	458	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2065	18.5	12.5	67.6	185	471	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2070	19.0	13.0	68.4	190	484	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2075	19.5	13.5	69.2	195	497	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2080	20.0	14.0	70.0	200	510	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2085	20.5	14.5	70.7	205	523	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2090	21.0	15.0	71.4	210	536	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2095	21.5	15.5	72.1	215	549	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2100	22.0	16.0	72.7	220	562	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2105	22.5	16.5	73.3	225	575	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2110	23.0	17.0	73.9	230	588	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2115	23.5	17.5	74.5	235	601	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2120	24.0	18.0	75.0	240	614	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2125	24.5	18.5	75.5	245	627	1.5	2.5	1.5	2.5	1.5	2.5
Algeria	2130	25.0	19.0	76.0	250	640	1.5	2.5	1.5	2.5	1.5	

FIGURE 118

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQEGEDPCVEAVGERGGPQNPD SRARLD
QSD EDFKPRIVPYYPDNPKNPYKKVLRTRYIQTELGSRRERLLVAVLT SRATLSTLAVAVNRTV
AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY
VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG
DILSÄRPDEWLGRCLIDSLGVGCVSQHQGQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE
GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLT VLTPEGEAGLSWPVGLPAPFTPHSRFEV
LGWDYFTEQHTFSCADGAPKCPLOGASRADVGDALETALEQLNRRYQPRLR FQKQRLNNGYR
RFD PARGMEYTLDDLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLP LL
VAEAAAAPAFLEAFAANVLEPREHALLTLLL VYGPREGGRGAPDPFLGVKAAA AELERRYPG
TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLT TVWTRPGPEVLNRCRMNAISGWQAFFP
VHFQEFNPALSPQRSPPGPPGAGDPDPPSPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA
RARLAGELAGQEEEEALEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR
CRLSNLEGLGGRAQLAMALFEQE QANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC
GTTCTAGACGCGGGAAAAATGCTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG
AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTA AAC
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTG
GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAA
AAAAGGATCCATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT
CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA
AAAGATGTATTTAATACCAAATCTGTTGGGCTTCTATTAAAGAGGCAATGACTTATCACCC
CAACCAGGTAGTAGAAGGCTGTTGTTTCAAGATATGGCTGTTACTTTTAATGGACTGACTCCAA
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTGCGGCATATTTTCAAT
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACATACATATCCAA
TACAGCTGTATGTTTCTTTTCTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG
TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTTCTTTAAACACATGAACATTGTAAATG
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTTGCAAATAAACTATTAATAAATATTATAT
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTGCTGATTGGTT
AAAAAATTTTAAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT
TGTGATTAAAGTAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACCTAAGCAACTAAAGAAA
ATTAAAGTGAAAGTTGAAAAAT

FIGURE 120

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME
LSKSFRVYCIILVKPKDVSLLWAAVKETWTKHCDKAEFFSSSENVKVFESINMDTNDMWLMMRK
AYKYAFDKYRDQYNWFFLARPTTFATIENTLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMIEGG
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFN
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL
PPNGSDND

Signal sequence:

amino acids 1-33

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,
399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,
558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG
AAAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAAATGAATTCATCTAAATCAT
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCAT
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTACAGAGCTCTCCTGCT
ACAATTATGGATCAGGTTCAAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC
AGCTGCTACTTCTTTTCTACTGACACCATTTCTGCGGCTTAAGTTTAAAGAACTGCTCAGC
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA
AACCTAAAATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG
AACAAAGGAAAATCTCTTTAAGAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC
ATAAGTATTTGTTACTCTGATACAAATAAAAAATAAGTAGTTTAAATGTTAAAAAAAAAAAAA
AAA
AAAAA

090964-071904
"02/20" 4366660

1000
900
800
700
600
500
400
300
200
100
0

Signal sequence:

N-glycosylation site.

Casein kinase II phosphorylation site.

N-myristoylation site.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

FIGURE 123

GGGACTACAAGCCGCGCGCTGCCGCTGGCCCCCTCAGCAACCCTCGACATGGCGCTGAGGCGGCCACCGCGAC
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACCTGCTTGCATCATTACGGATTTCG
AGACAAGTGACCCAGGATCGAGTGGAAAGAAAATTCAGATGAACAAACCATATGTGTTTTTTTGACAACAAAA
TTCAGGGAGACTTGGCGGGTCTGTCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG
ACTCAGCCCTTTATCGCTGTGAGGTCTGTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC
ACTGCCAGGAGAGTGAGGGCCACCCCGGCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACGGATT
CCAGAGCCAATCCAGATTTTCGCAATCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCCTGCTG
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTTATGGGGGGTTCGTTGTCTTGTCTGCTGACTGGCCCTGA
TCCGTTGGGCATCTGCTGTGCATACAGAGCTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA
ACCCAGGGAAACAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAAAGCTTTTCGTTTTGGCCAAAGTTGACCA
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCTCGCATGAGTATTAGG
GTGATCTTAAAGAGTTTGTCTACGTAACGCCCGTGTGCGGCCCTGTGAAGCCAGCATGTTTACCCTGGTCTGTT
CAGCAGCCACGACAGCACCATTGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCCGGCGGGAACCCA
GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAATCAGGAAG
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA
ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTGTTATTTTATAAAATTT
TACATCTAAATTTTGTAAAGGATGTATTTTGTATTGAAAAGAAAATTTCTATTTAACTGTAAATATATTGT
CATACAATGTTAAATAACCTATTTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT
TGGAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTTCTTTTCT
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCGCTCTAGCTCACTGTTGCTCGCTGTCTGCCAGGAGGCCCT
GCCATCCTTGGGCCCTGGCAGTGGCTGTGTCCCAGTGAGCTTTTACTCACGTGGCCCTGCTTTCATCCAGCACAGC
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAAACAGACCTCT
TTTTGGTTATGGATGGCTCACAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTGTTTAAATATTTGTT
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTAAAAAGAAAATGGAT
CCCCTGTTCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGCATTTCAAAACAAACCATGAT
GGAGTGGCGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG
TGAAACGCCTGAATCAAAGCAGTTTCTAATTTTGACTTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCTGGA
GAATGGCTCTCACTACTCACCTTGTCTTTTCACTTCCAGTGTCTTGGGTTTTTTTATACTTTGACAGCTTTTTTTT
AATTGCATACATGAGACTGTGTTGACTTTTTTTTAGTTATGTGAAACACTTTGCCGAGGCCGCTGGCAGAGGCA
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGTGTCATGGCATCCTGGATGCTTAGCATGCAAGTTC
CCTCCATCATTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACAGCTCCAGCCTCCT
TCTTGGTTGTATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA
TGGGAACAGGTCTGAAAAAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA
CGGAAAAGGAATACTCGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTAATAACACACCGTAATTTGGCATTTGTTTAA
CTCATTTATAAAAGCTTCAAAAAACCCA

FIGURE 124

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSWKIWNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDLNIIGGIIGG
VLVVLAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267